1978

An Analysis Of The Distribution Of Lifetime Purchasing Power In Canada

Ian Joseph Irvine

AN ANALYSIS OF THE DISTRIBUTION OF LIFETIME PURCHASING POWER IN CANADA

by

Ian Joseph Irvine

Department of Economics

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario

ABSTRACT

In this thesis, the distribution of lifetime purchasing power for a large sample of Canadian households is examined. The study differs from most distributional studies in two ways. First, it annuitises an estimate of life-cycle earnings and so is capable of encompassing variations in earnings which occur over the life-cycle and which are responsible for some of the observed inequality. Thus, it is dynamic in that it is not confined to examining the distribution of earnings or income at a point in time. Second, it focusses on the <u>joint</u> distribution of income and wealth in that the annuity distribution generated incorporates both life-cycle earnings and net assets. Thus, it omits neither stocks (wealth) nor flows (earnings/income) since the focus of attention is on purchasing power.

The degree of inequality associated with the distribution of annuities is found to be less than that associated with the more traditional measures of inequality, such as the distribution of wealth or income or earnings.

ACKNOWLEDGEMENTS

In completing this thesis I am very much in debt to my three advisors: Kul Bhatia, David Laidler (my Chairman) and David Scheffman. Each, at various stages, contributed to, and improved the quality of the final product immeasurably.

Many others at Western also made significant contributions. Tom Courchene and Tom Romer were responsible initially, with other members of my Committee, for getting me on the right track. Robin Carter was an endless source of ideas concerning measurement problems. Peter Chinloy and Walter Haessel read parts of the thesis and made valuable recommendations which were subsequently included. Statistics Canada supplied me with valuable data on net worth. I am grateful to them and to those in Western who facilitated this.

I would like also to thank Mary Clemente and Yvette Parent who typed early drafts of the thesis and particularly Joanne Lemon who patiently typed the final draft under deadline pressure.

Finally, I am most of all grateful to my parents who supported me in every way through my education. My debt to them is the greatest.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

In recent years there has been a renewed interest in many economies in the distribution of incomes across families and households. This revival has occurred despite the belief formerly held by some economists and sociologists that when the economic pie reached a certain size interest in its division would fade away. The fact that interest has not faded, despite the exponential increase in incomes over the last few decades, serves to focus our attention first on the view that poverty must be considered to be a relative rather than an absolute concept and second on the nature of the social structure in which economic activity takes place.

Most Western societies are in a certain sense split level structures. On the one hand, the socio-political institutions provide universal rights which imply the equality of all their members. On the other hand, economic institutions to a large extent rely on market forces to achieve their goal of efficient production and the result of this is that disparities in living standards are engendered.

In an economy dedicated to free enterprise, public efforts to promote greater equality are looked upon as an infringement on the outcome of market processes and as a result decrease the level of efficiency in the economy - which is seen by the proponents of this view as the most worthwhile goal. But to the proponents of a purely egalitarian structure, the existence of a competitive market place appears as the negation of certain basic rights.

Much effort has been spent by economists and politicians in devising methods whereby some balance can be maintained between the objectives of equality and efficiency. This process requires some measure of the degree of inequality which exists in the economy, for one clearly cannot, or should not, design policies which seek to attain some desired balance unless the existing degree of inequality is known.

The standard approach in studying problems of distribution is to examine the size distribution of earned income or a somewhat broader definition of income which includes income from wealth, capital gains, taxes, subsidies etc. The Gini index and Lorenz curve have been used most commonly as measures of inequality, although, more recently, other measures have found favour.[1] With some exceptions[2] the thrust of most studies has been to focus on the distribution of income <u>at a point in time</u>. This approach has many shortcomings: Income, as conventionally defined, does not adequately measure the purchasing power at an individual's command. Moreover, income at a point in time will not provide a good index of consumption and economic welfare over the life-cycle, and it will lead to dubious comparisons among individuals at different points in their life-cycles. These problems are fairly well-known. They pertain to the definition of income, to differences in asset-income ratios of different age groups, and to differences in the composition of income and wealth. They have been examined at various stages by Simons (1938), Vickrey (1957), Due (1960), Paglin (1975) and many others.[3]

This study is based on the general postulate that consumption is a more appropriate indicator of welfare than income. (Indeed it is surprising that so few studies have adopted this standpoint given the general philosophy of the discipline that it is consumption, rather than income or production, which is the ultimate goal of economic activity.) It is also postulated that households maximize utility or welfare over the course of the life-cycle. Given these two postulates, the objective of this thesis is to examine the distribution of optimal consumption streams for Canadian households. Given that incomes vary considerably over the life-cycle and that the distribution of income differs substantially from the distribution of wealth, we should expect that a consideration of these factors together would indicate a degree of inequality in the economy different from that which conventional studies of the distribution of income or earnings at a point in time indicate.

The estimates shall be derived mostly from cross-section data collected by Statistics Canada for the years 1970 and 1971. Ideally longitudinal data over the course of a lifetime would yield the best estimates. But as this exists in no economy (and will not in the forseeable future) we use cross-section data.

We shall follow this plan: In the next chapter, recent Canadian research in the general area of distribution is reviewed. These studies contain the main results with which the conclusions will be compared. The theoretical framework is set out in Chapter 3. Chapters 4, 5 and 6 deal with the data and estimation. The results are examined in Chapter 7. A summary and conclusions are given in Chapter 8.

NOTES ON CHAPTER ONE

1. See for example Atkinson (1971).

2. Nordhaus (1973), Paglin (1974).

3. If the correct definition is that given by Hicks, Haig and Simons, then income generated by net worth must be considered. The omission of this is likely to be very serious for older families who may have low observed incomes but have a high stock of net worth. Likewise, observed income for young families may be low, but potential income may be high.

CHAPTER 2

REVIEW OF CANADIAN STUDIES

Distributional studies can be classified for convenience into (a) static studies and (b) dynamic studies. The former deal with the distribution of some measure of economic welfare at a point in time, whereas the latter deal with a measure which might include potential earnings, the annuitised value of wealth etc. The number of studies of the latter type, of which this is one, is relatively small, but increasing.

An exhaustive review of the above is not attempted here. Instead, the major papers in each of the above categories will be reviewed in order to focus attention, first on answers which researchers have already come up with, and second on areas which require further research. It is hoped that this procedure will place the current study in context. Data sources and difficulties associated with the use of different types of data will also be examined since, as will be shown in the main body of this thesis, these problems are as great as the theoretical ones.

2.1 Static Studies

The first attempt to measure the distribution of income and wealth using survey data was that of Podoluk (1974). The major problem in this area lies in the difficulty in accessing reliable data. Data from financial institutions, while being possibly the most reliable, are not very useful for distributional studies because they do not provide information relating components of wealth to total wealth or total income. Estate tax returns are generally filed only by the higher wealth groups and thus are of limited use if the objective is to study the whole range of incomes or wealth groups. This leaves survey data as the only real alternative. On the surface, survey data appear to be the best means of obtaining information, since the survey will cover the whole range of the population and can be designed in such a way that the desired information can be directly obtained.

However, when surveys which were not primarily intended to serve as a base for distributional studies are used for this purpose, there arise the obvious problems of insufficient information and unrepresentative sampling. The problem of insufficient information is one which we will return to later in this study. As far as representative sampling is concerned, it is well known that, if a random sample is taken from the population, this will in general yield insufficient information on the behaviour of the upper tail of the distribution due to the positive skew in the distribution of both earnings and wealth. Some surveys have recognized this explicitly and have deliberately oversampled the upper tail (see Projector and Weiss (1966)). Even if the design and sampling issues are anticipated, there remains the problem of response reliability. This is particularly a problem where information on personal wealth is being sought as a typical respondent will underestimate his wealth holdings either because he genuinely omits some assets or deliberately does not report holdings for fear that he may not be paying sufficient taxes or for some such motivation.[1]

Finally, there are conceptual problems such as the valuation of pension plan contributions and insurance policies. Frequently, such a valuation can only be made on the date of maturity.

Despite these problems, Podoluk has been able to draw some interesting conclusions on the distribution of wealth in Canada on the basis of survey data. She estimated the net worth of families by wealth class and compared this distribution with the distribution of gross incomes by income class. This information is given in columns 1 and 2 in Table 2.1.[2] It is clear that the distribution of wealth is much more unequal than the distribution of incomes since the Lorenz curve for the former lies everywhere outside that for the latter (see Figure 2.1).

A more interesting question than those answered in columns 1 and 2 would be: how are assets distributed across income groups? An answer of sorts to this is given in column

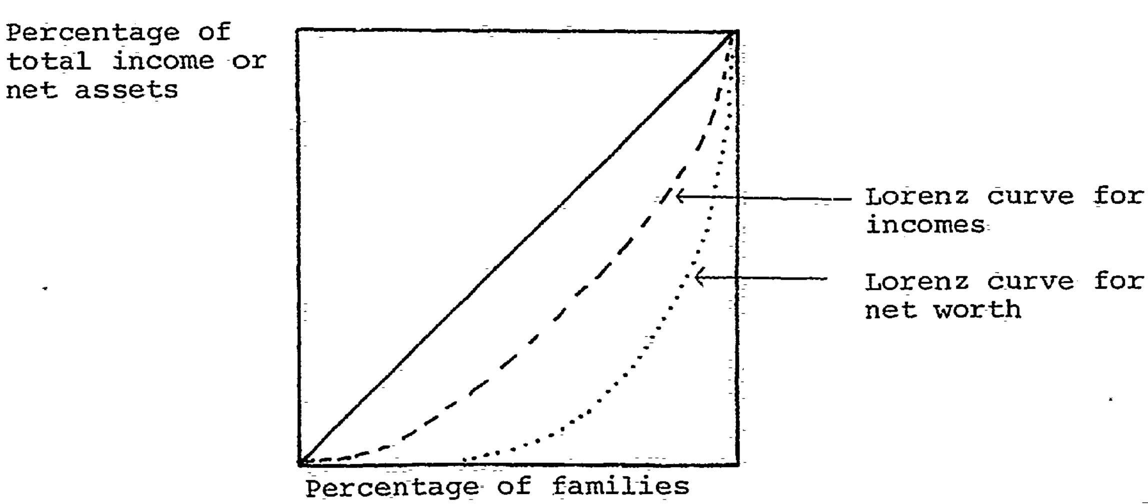
TABLE 2.1

DECILE SHARES OF INCOME AND ASSETS OF ALL FAMILY UNITS RANKED BY SIZE

	1969 Income	1970 Net Worth	Net Worth by Income Class
1st Decile	1.3	-0.9	3.7
2nd Decile	3.0	0.0	5.9
3rd Decile	4.7	0.2	6.9
4th Decile	6.4	1.1	6.7
5th Decile	8.0	2.8	7.7
6th Decile	9.6	5.2	7.5
7th Decile	11.3	8.2	9.4
8th Decile	13.3	11.8	11.1
9th Decile	16.2	17.7	13.2
10th Decile	26.2	53.9	27.9
Gini Index	.383	.724	.298

Source: Survey of Consumer Finances 1970 tabulated in Podoluk (1974)

FIGURE 2.1



3. This tells us that the lowest 10% of families ordered according to income own 3.7% of total assets, the lowest 20% own 9.8% etc.. The fact that assets are more evenly distributed between income groups than is income should not be a big surprise given that many of the households in the lower income groups may be retired and are at a stage in the life-cycle where home ownership is high and a large stock of durables has been accumulated.

These data, however interesting, do not provide a picture of how potential purchasing power is distributed even in a static framework because they do not take into account the redistributive effects of government activity. Two major studies have recently been carried out in Canada within a more general framework (Dodge 1975, Gillespie 1976) which attempted to assess the effects of government expenditures and taxation on the distribution of personal income in Canada.[3]

Dodge estimated the impact incidence of federal, provincial and local government taxes and expenditures for 1970.

Broad income[4] was first calculated for a sample of families and individuals on the basis of data from the 1969 Survey of Consumer Finances. Indirect taxes and expenditures by government on goods and services were allocated on the basis of family expenditure patterns estimated from the 1969 Family Expenditure Survey. Direct taxes and transfers were estimated on the basis of income and demographic characteristics of the family. The sum of total benefits and broad income minus total tax payments yields adjusted broad income. The difference between broad income and adjusted broad income is 'net fiscal incidence'. It is the distribution of this incidence across income groups in which Dodge was primarily interested.

In estimating the incidence, assumptions must be made about the distribution of expenditures across income classes and the degree to which taxes are shifted. The benefits from pure public goods were assumed to be distributed on the basis of broad income. The distribution of benefits from direct expenditures poses fewer problems than that for public goods. Dodge points out that expenditure incidence is sensitive to the assumptions made about the distribution of benefits from public goods but the conclusion that net incidence is broadly redistributive holds under a wide range of assumptions. Likewise, the estimates of net fiscal incidence were not found to be sensitive to assumptions made about the shifting of corporation and payroll taxes.[5]

Dodge then estimated the effects of taxation and the transfer programmes across fifteen money income classes. The impact of programmes at all levels of government was to provide large net benefits to families and individuals with incomes of less than $4,000, to provide declining net benefits to families earning between $4,000 and $11,000 and to levy a small but increasing net tax on families with incomes in excess of $11,000. Varying these assumptions by substantial amounts did not lead to significant changes in the results.

Gillespie's results (1976) are similar to those of Dodge. He concluded that,

> ...the total public sector in Canada was broadly redistributive from higher income to lower income classes in 1969. In addition, the public sector was redistributive in such a way that the gainers with higher incomes gained relatively less from, and the net contributors with higher incomes contributed relatively more to, the redistributive mechanism of the public sector.

The second objective of Gillespie's research was to examine the redistributive effects of government activity over time, with particular reference to the sixties. A surprising conclusion was that the increasing contributions of middle and upper middle income families to the redistributive mechanism of the public sector generated greater gains in absolute terms for the rich than for the poor. This phenomenon may be attributable to the general principle which has guided social security legislation in Canada in recent decades i.e. universality of coverage. Podoluk (1970) discusses this issue in some detail.

2.2 Dynamic Studies

While these studies answer many of the questions posed by the existence of a government sector, they do not answer questions posed by the dynamic behaviour of incomes over time or the effect of asset ownership on purchasing power. However, a small but increasing number of researchers are focussing attention on these problems. This development in distributional studies is not just due to the fact that the theoretical issues have received more attention in the last decade, but is equally attributable to the increasing existence of and accessibility to micro data.

Wolfson (1977), using micro data from the 1970 Survey of Consumer Finances, examined how the distribution of income changes when three adjustments are made to observed income. These are: adjustments for family size, the inclusion of imputed rent from owner-cccupied houses and the inclusion of the annuity equivalent of net worth. Using a methodology developed by Weisbrod and Hansen (1968) he estimated the annuity by the formula

$$f(n,r) = r/(1-(1 + r)^{-n})$$

where n = life expectancy

r = interest rate

$f(n,r)$ = income stream that can be purchased for \$1

Clearly, older families can purchase a greater annuity than younger families with a given wealth stock.

The effect of the inclusion of the annuity, imputed rent and family size considerations on the distributional measures is important to the extent that it increases the 'income' which accrues to older family units. This is because family size is small, house ownership rates are high and the net asset stock is high for this group relative to the rest of the population. Wolfson found that the addition of imputed rent to income reduced inequality. This initially may seem surprising given that the home ownership distribution is more unequal than the earnings distribution. He suggests that the answer must lie in the joint distribution: e.g. low income rural families tend to have high home ownership rates.

In contrast to this, and also to the results obtained by Weisbrod and Hansen, he found that the inclusion of the annuity equivalent of net worth did not change the degree of inequality substantially. The Gini index fell by 2.7% using a four percent rate of interest and rose by 2.2% using a ten percent rate.

A recent study by Henderson and Rowley (1977) may be considered to be a member of the class of dynamic studies to the extent that it decomposes the income distribution both at a point in time and through time and it examines the influence which economic and demographic factors - which change over time - have on the distribution of income.

Henderson and Rowley (H.R.) initially point out that the degree of inequality in the distribution of family incomes, as measured by the Gini index, decreased between 1951 and 1965 but increased by an approximately similar amount between 1965 and 1971. Their objective is to examine the causes for the changes which took place during the latter time period.

To answer this question, H.R. start by defining the evolution of Gini coefficients for various sub-populations. They show that the structure of the Canadian population changed substantially over the period. Average family size decreased from 3.4 persons in 1965 to 2.94 persons in 1973. In addition, a larger proportion of family units had heads of age 35 or less in recent years. The proportion in 1965 was 27% and in 1973 it was 32.5%. With a larger proportion of families headed by younger people and a larger proportion of single person families, it is to be expected that inequality in family incomes would increase. This is because young workers typically have lower earnings than older workers and with an increasing number of young workers breaking away from their (parent's) family and starting their own we observe a greater concentration of families in the lower tail of the distribution. The apparent increase in inequality can thus be attributed to changes in the demographic structure rather than to an increase in inequality among families of a given type.

H.R. further examine the inequality which exists among groups of given family size, of given education of head, of given number of earners etc. They then proceed to examine average incomes, proportional distributions and growth rates in incomes for various socio-economic groups. By standardising for changes in these socio-economic characteristics they standardise the summary measures of inequality and are able to estimate what proportion of the change in inequality is attributable to each of these factors.

The results of the H.R. study are interesting, first because they throw new light on the subject in Canada and second because they accord with the results obtained in similar studies using United States data (Rivlin 1975, Henle 1972). In the U.S., the distribution of incomes as measured by the Gini index has not changed much since the early fifties. But as in Canada, this is not because the determinants of the distribution have remained stable rather it is because they have acted in such a manner as to offset each other. There is thus no reason to suppose that because of the apparent stability in the distribution over the last two decades that this will continue to be the case in the future.

This recent work has served to focus attention on the fact that studies which use a static framework suffer from having no theoretical understanding of the processes whereby the behaviour of the incomes they purport to study are generated. Thus, they must be viewed purely as a description of the configuration of incomes, which change in systematic and non-systematic fashion over time, at one particular instant. In order to develop a measure of inequality which takes into consideration these dynamic essentials, it is necessary to develop an understanding of the manner in which incomes are generated.

In Chapter 3, such a framework is developed. We start with a brief examination of the historical development of distributional models. We then proceed to examine optimal lifetime behaviour on the part of households and show that this leads logically to the objective defined in Chapter 1: That is, distributional studies should ideally consider lifetime consumption streams.

NOTES ON CHAPTER TWO

1. Davies (1978) has discussed this issue in some detail and provides estimates of the degree of underestimation of net worth.

2. Column 1 states that the lowest 10% of families (ranked by income) had only 1.3% of total income. The lowest 20% of families had 4.3% of total income and so forth. Column 2 gives the same information except that families are here ranked according to net worth rather than income.

3. Gillespie did a similar type of analysis in 1966 for the Royal Commission on Taxation.

4. Broad income is defined as the sum of money income plus non cash income less government transfer payments and bond interest.

5. The precise assumptions underlying Dodge's research are the following:

   (i) 25% of corporation income taxes are shifted forward.

   (ii) The employer's share of social security taxes is borne totally by the employees.

   (iii) Estate taxes are borne entirely by those with incomes over $15,000. Indirect taxes are borne entirely by consumers. Income taxes are borne entirely by the income recipient.

   (iv) Costs of government supplied private goods are distributed on the basis of direct consumption while expenditures on public goods are allocated on the basis of broad income.

   (v) The effects of government deficits are ignored.

CHAPTER 3

THEORY AND MEASUREMENT

3.1 Historical Development

The earliest models which sought to explain the distribution of income strangely had very little economic content. Most took as their starting point some observed form such as the Pareto or Lognormal distribution and sought to develop a stochastic process which, operating within the economic system, would produce a stationary income distribution approximating that observed.

Kapteyn (1903) and Gibrat (1957) were the first to develop a 'law of proportionate effect' in a Markov model. The motivation for this sprang from the belief that the effects of many random factors causing changes in the value of a variate (e.g. income) over time are likely to be proportional to the value of the variate in the preceding time period. Thus, if income in any time period is specified to depend on the previous period's income, it can be shown (by using the multiplicative central limit theorem) that the Markov process generates a distribution of incomes which tends toward the lognormal.

The other distribution most frequently used in this literature is the Pareto, so named because Pareto (1897) noted a remarkable consistency in the distribution of incomes across different economies.

The original models of Gibrat and Pareto have been developed considerably by Champernowne (1953), Solow (1951), Aitchison and Brown (1954), Rutherford (1955), Mandelbrot (1961), Wold and Whittle (1957), Friedman (1953) and many others.

However, to consider a stochastic mechanism as the sole determinant of the distribution of income and to disregard the theory of choice completely is surely to abandon any hope of an economic or rational explanation. Blinder[1] sums the situation up in the following terms.

> It (the stochastic process model) is antithetical to the main stream of economic theory which seeks to explain complex phenomena as the end result of deliberate choices by decision makers. ..... one may think of the deterministic part of any model as 'what we (think we) know' and the stochastic disturbance as a measure of our ignorance. The probabilistic approach to distribution theory appears to allocate the entire variance in income to the latter. One would hope that economics could do better than that.

A second group of economic models which could broadly be termed 'ability-earnings' models represent a step forward from the stochastic process models to the extent that they incorporate economic factors into the processes which determine the form of the distribution of income.

Pigou (1932) proposed that the reason why a skewed distribution of incomes exists is that there exist non competing groups in the labour market. Miller (1955) and Staehle (1943) provided evidence to support Pigou's theory. Some useful theorems derived by Haldane (1942) provided encouragement to economists to relate the distribution of abilities to the distribution of incomes and since that time much work has been done by Roy (1950a, 1950b, 1951), Mayer (1960), Lydall (1959, 1968) and others in an attempt to develop an analytical form for the distribution of incomes which would take into account factors such as ability, responsibility, number of employees etc..

Tinbergen, in a series of papers (1951, 1956, 1957, 1970) proposed a model in which the distribution of income is the result of the interaction of a demand for labour function and a supply of labour function where the latter is based on utility rather than income maximization. If jobs require certain abilities and individuals possess other abilities, this system introduces tensions between attributes required and attributes available. If individuals have a quadratic utility function, Tinbergen shows that this process can generate the observed distribution of incomes.

These models represent an improvement over the purely stochastic models. First, they distinguish between earnings and wage rates. Second, there is the recognition that ability is an important factor. However, to the extent that the implications depend on a specific form of the distribution of abilities, the models are not productive since Mayer (1960) has shown that we can always devise intelligence or ability tests to yield any form of distribution we wish.

The major shortcoming however is the neglect of the theory of choice and the focussing of attention on an effort to develop a mechanistic theory which would explain the existence of a particular mathematical form of the distribution.

Inherent in the models described so far is the implication that, for the most part, the distribution of earnings is not affected by individual choice. In contrast, the human capital models focus on individual investment behaviour as a determinant of the distribution of earnings. Mincer (1958) was the first to examine earnings distributions from this standpoint. He developed what has become known as the 'schooling model'. In this initial development, he assumed equal opportunities and abilities existed and examined the properties of the resulting earnings distribution if individuals have the same present value of lifetime earnings. The purpose of making these assumptions (which have since been relaxed) was to enable an examination of the effects of differential training on labour incomes in an environment free from non-competitive forces.

This simple model is characterized by a remarkable number of properties which are observed in empirical studies. It can be shown that people with more schooling earn more income, that the difference in earnings increases with the rate of return, that the difference in earnings is inversely related to the working life span and that relative differences in income are greater at higher schooling levels even if the difference in years of schooling remains the same. This model also predicts that (a) the distribution of income is log-normally distributed and thus skewed, (b) relative dispersion in earnings is larger the larger the dispersion in schooling, and (c) relative skewness increases with the dispersion in the distribution of schooling. This model, with some modifications, has been tested by Chiswick (1974) and Mincer (1958). Becker (1964) and Ben-Porath (1967) examined the question of how investment in human capital should be made over the life-cycle and many other studies in the last decade have extended the basic model.

The shortcomings associated with the human capital model are fewer than those associated with the models examined previously. Apart from the measurement of human capital, the major problem lies in the assumption that individuals seek to maximize the present value of lifetime earnings rather than some function of consumption and leisure. In a sense, the human capital model could be said to be applicable to the distribution of potential earnings but instead is frequently used as a description of the actual earnings distribution.

The model to be used in this study is more in the spirit of Nordhaus (1973) and Blinder (1974). In particular, it is based on explicit utility maximization rather than earnings maximization. It also integrates labour and non-labour income into one quantifiable measure. We now examine this model in detail.

3.2 Optimal Household Life-cycle Behaviour

We first consider the case of a single individual maximizing lifetime utility. We will then extend the model to incorporate the effects of family size and finally define the estimation of an annuity for each household.

Consider an individual with an instantaneous utility function $U[C(t)]$ who maximizes lifetime utility over a known horizon T. Then

$$U = \int_0^T U[C(t)]e^{-\delta t}\, dt \qquad (3.1)$$

where δ is the subjective rate of discount.[2] This is maximized subject to lifetime wealth W_o where this latter is composed of the net stock of assets K_o and the present value of lifetime earnings. The constraint is thus

$$W_o = K_o + \int_0^T Y(t)e^{-rt}\,dt \tag{3.2}$$

where $Y(t)$ is labour income in any time period and r is the interest rate facing the consumer. Time period zero is the year in which the individual enters the labour force or leaves his parents' home to start his own household, i.e. the year in which he becomes a 'homo economicus'. Up to this time period we assume that all decisions have been made for him by his parents. This is an important assumption because it means that educational choices are not a part of the decision-making process of the homo economicus. Each person is thus envisaged as entering the labour force with a given amount of human capital which defines in part his future earnings.

A solution for (3.1) and (3.2) is obtained by maximizing (3.1) subject to the constraints that consumption plus saving equal total income at each instant

$$C(t) + S(t) = Y(t) + rK(t) \tag{3.3}$$

and that the change in the stock of wealth equals saving at each instant

$$\dot{K}(t) = S(t) \tag{3.4}$$

where $\dot{K}(t)$ is the time derivative of $K(t)$.

The Hamiltonian function is defined by

$$H(K,S) = U[Y + rK - S]e^{-\delta t} + \mu S \tag{3.5}$$

where μ is the costate, S the control and K the state variable. Time subscripts are omitted for convenience. The conditions

$$\frac{\partial H}{\partial S} = 0 \quad ; \quad \frac{\partial H}{\partial K} = -\frac{\partial \mu}{\partial t} \tag{3.6}$$

are necessary and sufficient for a maximum if U(C) is concave. Conditions (3.6) yield

$$-U'(C)e^{-\delta t} + \mu = 0 \tag{3.7}$$

$$U'(C)e^{-\delta t}\, r = -\dot{\mu} \tag{3.8}$$

Differentiating (3.7) with respect to time and substituting into (3.8) yields a solution for the time path of consumption.

$$\dot{C} = -\frac{U'}{U''}(r - \delta) \tag{3.9}$$

Since U" is negative then $\dot{C}$ is positive if $r > \delta$. This means that if the interest rate exceeds the rate of time preference consumption is increasing over time since the individual would prefer to invest capital and consume later. Conversely if the rate of time preference exceeds the interest rate consumption is decreasing.

At this point, it is necessary to specify an exact form for U(C) if we are to be able to do more than make qualitative statements about consumer behaviour. The function which will be utilized here is the isoelastic form

$$U(C) = A\,C^{\beta}/\beta \qquad \big|\, -\infty<\beta<1\ ,\ \beta \neq 0$$
$$U(C) = A \log C \qquad \big|\, \beta = 0 \tag{3.10}$$

Yaari (1964) has pointed out that if we believe, in response to a given increase in lifetime wealth, the optimal $C(t)$ should increase in the same proportion for all t, the utility function must be isoelastic. Blinder (1974) notes that the function also displays constant relative risk aversion.

Using this function, equation (3.9) becomes

$$\frac{\dot{C}}{C} = \frac{(r-\delta)}{(1-\beta)} \tag{3.9a}$$

which has as a solution

$$C(t) = C_o e^{gt} \tag{3.9b}$$

where $g = (r-\delta)/(1-\beta)$ and C_o is consumption in the base year. To obtain a numerical solution for $C(t)$ it is necessary to express the initial consumption C_o in terms of lifetime wealth. If all wealth is consumed during the lifetime it follows that

$$W_o = \int_0^T C(t)e^{-rt}dt = \int_0^T C_o e^{(g-r)t}dt \tag{3.11}$$

Integrating (3.11) and solving for C_o yields

$$C_o = \frac{(r-g)W_o}{1-\exp[(g-r)T]} \tag{3.12}$$

which, when substituted into the optimal value for $C(t)$, yields

$$C(t) = \frac{(r-g)W_o \exp(gt)}{1-\exp[(g-r)T]} \tag{3.13}$$

At this point, we can see that it is possible to estimate the value of consumption in every time period which would maximize lifetime utility assuming that the parameters in (3.13) are known and that an estimate of W_o can be obtained.

The above results however are not sufficient to describe the optimal behaviour of multiperson households because they pertain to an individual. The fundamental question when dealing with household behaviour is how to modify the objective function (3.1) to incorporate the effects of family size. One approach is to specify utility as linear in the number of persons times the utility of consumption per capita.

$$U = \int_0^T F_t \, U(\frac{C_t}{F_t}) e^{-\delta t} \, dt \tag{3.14}$$

Another is to specify utility as a function of per capita consumption.

$$U = \int_0^T U \, (\frac{C_t}{F_t}) e^{-\delta t} \, dt \tag{3.15}$$

There are problems associated with both of these formulations. Arrow and Kurz (1970) point out that the use of (3.15) implies that we should allocate <u>less</u> consumption per capita to more populous generations. On the other hand, the use of (3.14) could yield the result that utility for a large family with low per capita consumption could exceed that of a small family with high per capita consumption.[3] We will examine this problem again at the estimation stage.

3.3 A Distributional Measure

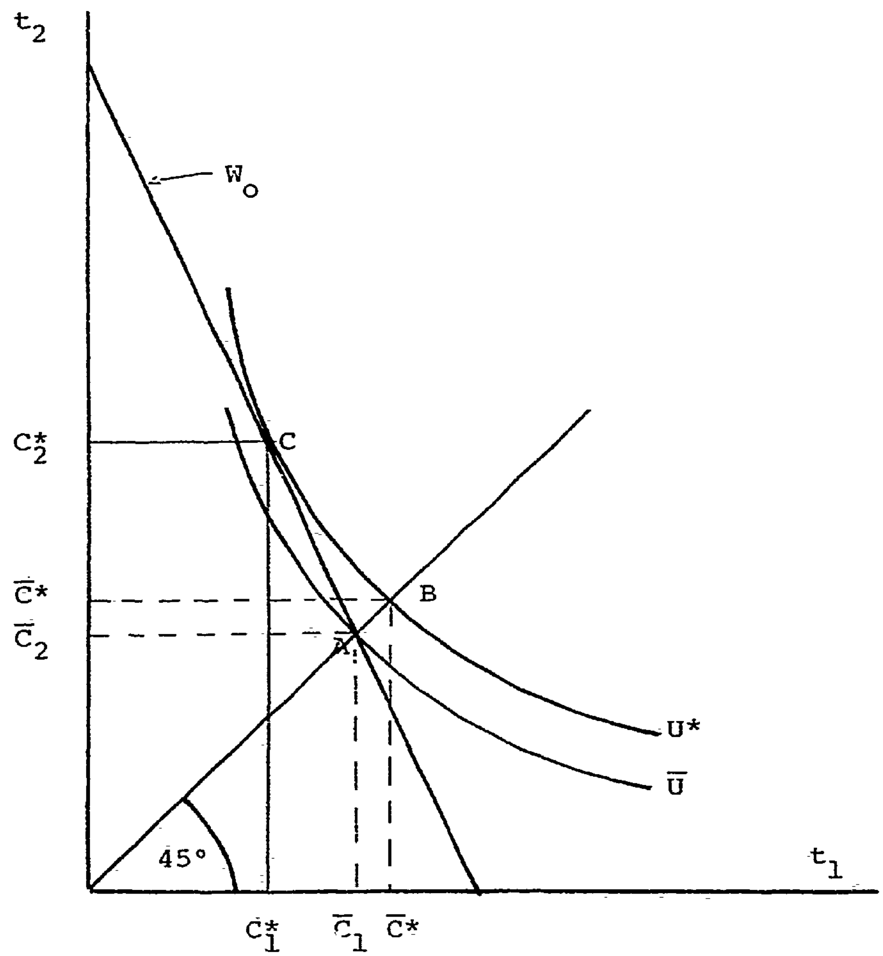
We now examine the possibility of developing a distributional measure which will enable us to compare people with different asset income ratios, different life expectancies and different lifetime earnings.[4] Since we envisage families as utility maximizers, we focus on a distributional measure which is based on the utility per unit of time derived by a family from consumption. Since we assume the family knows its lifetime wealth, it also knows its optimal consumption path. In general, the amount of consumption per unit of time is not constant. However, it is possible to estimate an annuity which would yield the same lifetime utility as the optimal consumption path. Such an annuity is the distributional measure we will use.

Consider Figure 3.1 where the concept is illustrated for a two period world.

Consider an individual who lives for two periods whose present and future incomes are known with certainty. If, in addition, he knows his current wealth position, his lifetime budget constraint can be estimated. In Figure 3.1, this lifetime wealth is denoted by W_o and the slope of the budget constraint is given by $-(1+r)$, where r is the real rate of

FIGURE 3.1

UTILITY EQUIVALENT ANNUITIES



return on wealth holdings. Given an inter-temporal preference set, this level of wealth permits the individual to attain an indifference set defined by U^* and his optimal consumption plan is defined by $\{C_1^*, C_2^*\}$. Now, if he had constant annual consumption and was subject to the same budget constraint, he would attain a lower level of indifference $(\bar{U})$, unless his rate of time preference was equal to the rate of return in which case point $A\{\bar{C}_1 = \bar{C}_2\}$ would coincide with point C. However, if his equal consumption in each year was defined by $\{\bar{C}^*, \bar{C}^*\}$ this would mean he would be indifferent between constant consumption in each period and the optimal consumption plan $\{C_1^*, C_2^*\}$ since both leave him on the same level of indifference.

The objective then is to estimate a distribution of $\bar{C}^*$'s across households and then analyse how this distribution (which we will term the annuity distribution) differs from the distribution of observed incomes at any point in time.

A numerical solution for the annuity (which is termed the 'utility equivalent annuity income') is obtained by solving

$$\int_0^T \frac{A}{\beta}(\bar{C}^*)^{\beta} e^{-\delta t} dt = \int_0^T \frac{A}{\beta}(C^*(t))^{\beta} e^{-\delta t} dt \qquad (3.16)$$

for $\bar{C}^*$, which is the annuity. $C^*(t)$ is the solution for the optimal consumption path. Integrating both sides of (3.16) and solving for $\bar{C}^*$ yields

$$\overline{C}* = \frac{(r-g)\ W_o}{1-\exp[(g-r)T]} \cdot \{\frac{\delta[1-\exp[g\beta-\delta]T]}{[1-\exp(-\delta T)](\delta-g\beta)}\}^{1/\beta} \qquad (3.17)$$

The solution for the annuity differs from the above in the case of multiperson households. The explicit solution is postponed until a later stage as is the estimation of the various constants in (3.17).[5]

In the following chapters, we shall estimate a value of $\overline{C}*$ for a sample of families. The resulting distribution of annuities will then be compared with the results of studies outlined in Chapter 2. The focus will be on the estimation of W_o since this, with lifespan and family size, are the values in (3.17) which vary across households. [We will also experiment with varying some of the constants at a subsequent stage.] The components of W_o are lifetime earnings of the family and the net stock of assets in the base year. The estimation conforms to the following pattern.

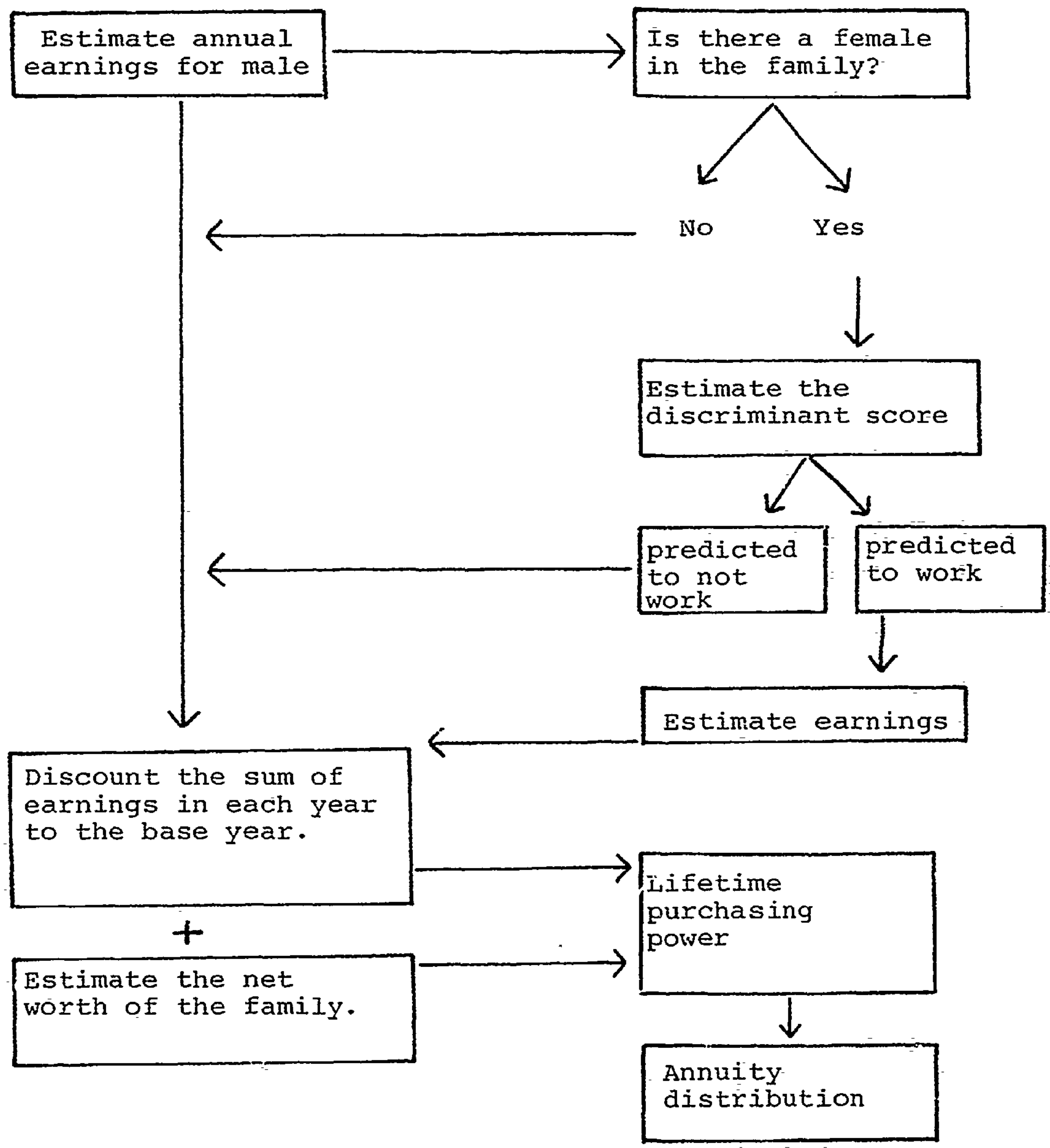
Each economic family earns a certain income in each time period until the death of both parents, or the death of the head for families where there are not two parents. The male is assumed to work in each time period and the female's work decision (if there is a female in the family) depends on such factors as the male's income, number and age of children, age etc.. The work/no work decision is the outcome of a discriminant analysis. If she is predicted not to work then the family's income for that year is the income of the male.

The predicted earnings are obtained from our knowledge of the economic and demographic characteristics of the individual. The lifetime incomes are then discounted back to the base year. The sum of discounted incomes and net worth of the family represents lifetime purchasing power.

Appendix A describes the data base in detail. Appendix B contains a discussion of the households included in the estimation. Chapter 4 deals with the estimation of earnings for males. Chapter 5 deals with the same issue for females. In Chapter 6, we estimate the net worth of the families and examine some issues pertaining to the estimation of the annuities. Figure 3.2 summarizes the procedure.

FIGURE 3.2

ESTIMATION PATTERN OF LIFETIME PURCHASING POWER



NOTES ON CHAPTER THREE

1. Blinder (1974) p. 9.

2. If T is not known with certainty but instead follows a known probability distribution, the results do not change substantially. See Appendix E on this.

3. For a detailed discussion of this point, see Arrow and Kurz (1970) p. 14.

4. This section draws on some results of Nordhaus (1973).

5. It is worth pointing out that the above is purely a purchasing power model. That is, the utility function does not attribute a financial value to leisure.

   It could be argued that if individuals are free to choose their hours of work then the utility obtained from work and income should be equal at the margin and thus a value should be imputed to leisure based on the wage rate. This procedure has not been adopted here (a) because the required assumption of flexible hours seldom holds in practice and (b) because of the difficulty associated with predicting a wage rate for females over the course of their life-cycle from cross-section data. Nonetheless, it is true that the measure of economic well being used in this thesis falls short of a more comprehensive definition which would impute a value to leisure and non-market activity.

CHAPTER 4

THE ESTIMATION OF LIFE-CYCLE EARNINGS FOR MALES

In this chapter, we derive estimates of lifetime earnings for a sample of male household heads in Canada. We proceed as follows: The first section deals with a methodological problem relating to the use of cross-section data in predicting earnings behaviour over time. The second deals with the specification of earnings equations. In the third section, the data and estimation problems are outlined. In the final section, we present the results.

4.1 Methodological Issues

The data used in this study have been obtained from a cross-section. Thus, there arises the fundamental issue of how to extrapolate future earnings patterns using regression coefficients which have been estimated for one point in time.

The relationship between age and earnings at a particular time has been well documented by Hanoch (1967), Becker (1964), Mincer (1974), Friedman and Kuznets (1954) and others. Earnings first rise quite steeply in the early years, reach a plateau about the age of fifty and tail off before the age of retirement. This behaviour is shown in Figure 4.1.

Using results based on the relationship indicated in

FIGURE 4.1

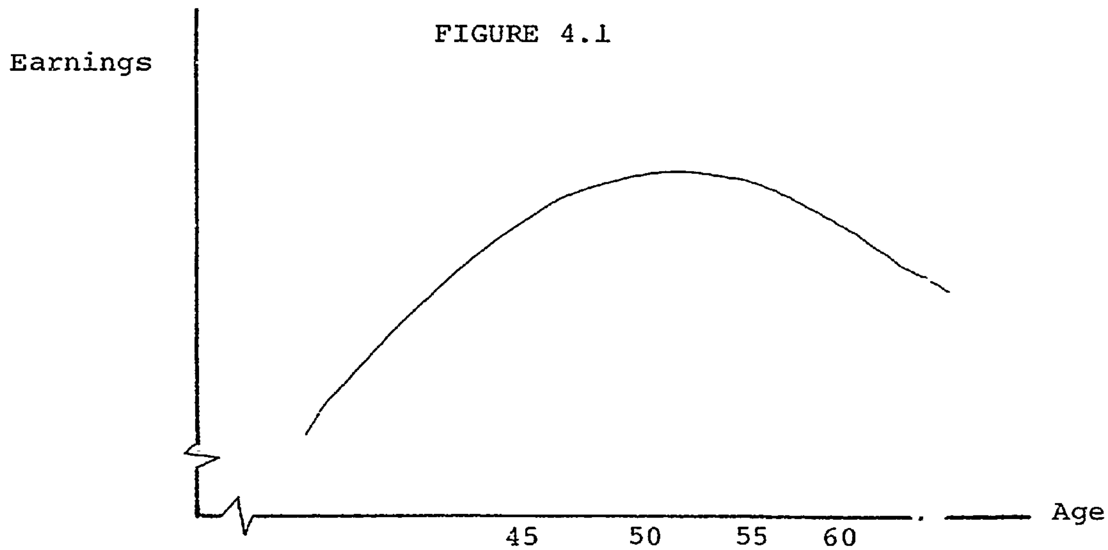
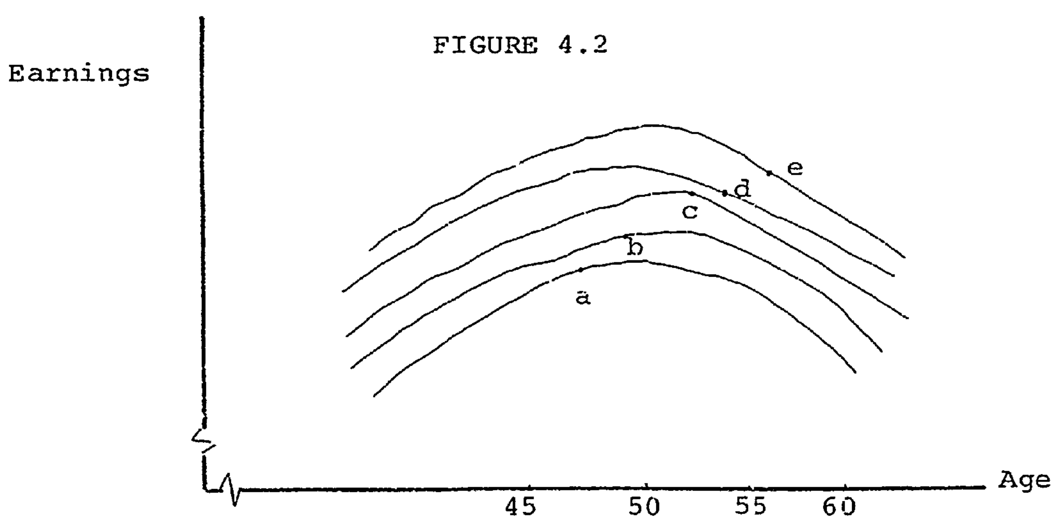


FIGURE 4.2



Hypothethical Age - Earnings Profiles

Figure 4.1 to derive earnings profiles would thus involve imputing declining earnings to individuals once they reach age 50. But this is in conflict with findings based on longitudinal and cohort analysis which indicate that earnings continue to increase (though at a decreasing rate) until the age of retirement.

In a recent study, Ruggles and Ruggles (1974) examined the seemingly conflicting evidence from cross-section and time series studies using a cross-section longitudinal data base. This base was a one percent sample of (U.S.) social security files which were observed over the time period 1957-1969. Their findings indicated that there was a continual (though not identical) upward shift in the age-earnings relationship over the period. This shift is probably attributable to secular growth in the economy.

Thus, when an individual is observed over time, he is really moving from one cross-section profile to another and thus his earnings need never decrease. This is shown in Figure 4.2 where a hypothetical range of the profile is drawn for several years. In the first year (age 47), he seems to be at the peak of the cross-section profile but earnings continue to rise along the path a, b ....e.

The problem at the estimation stage is that no such data exist for Canada and thus some method of approximating the increase over time must be used. Fortunately, there is evidence to indicate that quite simple adjustments to the cross-section regression results yield reliable time series earnings profiles. Taubman and Wales (1974) used cohort data for the period 1946-1969 to estimate the returns to education. Using the observed data for the whole period ex post returns were estimated and these were then compared to returns derived ex ante by predicting earnings profiles on the basis of cross-section data for 1949 after allowing for technical change.

Their correction procedure is to boost predicted earnings by a certain percentage in each year. The close similarity of the results obtained from each approach (i.e. ex post and ex ante) indicate that their method of correcting the cross-section data to allow for economic growth provides a good approximation to the real earnings profiles. This is the approach that is adopted in correcting for the difference between the cross-section and the time-series.

4.2 The Specification of Earnings Equations

There has been a considerable amount of research over the last decade on the estimation of earnings and income functions. Much of this has sprung from an interest in estimating the returns to education and as a consequence many of the models have been of the human capital type.[2] These have focused primarily on the relationship between years of schooling and earnings. By and large, these studies have used cross-section data, but an increasing number have used longitudinal or panel data with the improving though still limited availability of this type of data.[3]

Systems built on this basis have had some degree of success in explaining observed patterns of earnings and the resulting distribution of these earnings. More recent work[4] has sought to extend the conventional human capital models to incorporate the effects of post-schooling investments in human capital, i.e. "on the job" training. Recent research[5] has also focused on the role of education as a screening device and the relationship between this and the optimal expenditure on education.

As pointed out in Chapter 3, we view the distribution of earnings as being primarily the result of a given distribution of human capital. The purpose of this section is to examine how to specify earnings equations based on the human capital model which will enable us to project earnings into the future for the sample of families.

The starting point for the human capital model is to assume that the present value of lifetime earnings from different jobs or professions is approximately the same. If this is the case then the present value of earnings of two individuals - one with s years of schooling and the other with s-d years - can be formulated respectively as

$$
\begin{aligned}
V_s &= Y_s \int_s^n e^{-rt}\, dt \\
V_{s-d} &= Y_{s-d} \int_{s-d}^n e^{-rt}\, dt
\end{aligned}
\tag{4.1}
$$

where Y_s and Y_{s-d} are the annual incomes, r the discount rate and n the year of retirement.

Equilibrium in a competitive human capital market requires $V_s = V_{s-d}$. Using this condition and integrating both sides of (4.1) we obtain

$$
Y_s/Y_{s-d} = (e^{r(n+d-s)}-1)/(e^{r(n-s)}-1) \tag{4.2}
$$

and if n is large, this becomes $Y_s/Y_{s-d} = e^{rd}$ [6]. Mincer (1974) points out that the above formulation has properties which correspond with those which have been observed in empirical studies. These properties have been outlined in Chapter 3, section 3.1. However, he also showed that the model does not provide a really adequate explanation of earnings patterns.

Using as sample data white nonstudent nonfarm males between the ages of nineteen and sixty-four, he was able to explain only 7% of the variation in earnings. A preliminary run on Canadian data indicates an explanatory power of approximately the same magnitude as Mincer found.[7] This does not mean that the schooling model is irrelevant but does indicate it is insufficient. There is good reason to suspect the model's performance would be improved if post-school training could be incorporated. For example, it has been observed[8] that a twenty-five year old who has completed school at the age of twenty-two frequently earns less than a person of similar age who left school at seventeen. Thus, even though he has more years of formal education, his income is still lower because of fewer years experience; i.e. his post-school human capital investment is not abreast of that of the younger school leaver and this is reflected in his earnings.

The effect of this on differential earnings will be more noticeable among the relatively young members of the work force as the individual with more years of formal schooling, while initially earning less than the amount earned by the individual with fewer years schooling but more experience, will generally overtake the latter within seven or eight years of completing school.[9] The problem then is how to incorporate the experience or "on the job training" variable into the model.

4.2.1 The Learning Model

Assume that earnings in any given time period t are defined as some function of accumulated human capital (H), a set of shift parameters (D) and a random component (ε).

$$Y_t = Y_t(H, D, \varepsilon) \tag{4.3}$$

Human capital is now defined as a function of years of schooling (S) and ability (A) - which depends on years of experience (x). For estimation purposes, x can be considered to be present age minus age at which the labour force was entered and S the age at graduation minus five.

The time path of the individual's ability is dependent upon the change in efficiency with which he performs his work. This formulation, we shall see, enables us to bypass the problem of defining such a nebulous variable as 'ability'. Following Eckaus et al. (1974) efficiency is specified in accordance with learning theory, which suggests that the rate of change of the individual's ability or efficiency is inversely related to the number of times which the task has been performed. Thus, by treating each year as the equivalent of the completion of one task, we can formulate the change in the individual's efficiency as

$$\frac{\frac{dA}{A}}{dx} = \frac{a}{x} \tag{4.4}$$

where a is some constant. When integrated, this yields

$$\log A = a \log x + k \tag{4.5}$$

where k is the constant of integration.

A solution for k can be obtained by letting the number of years of experience (x) equal 1. Thus, (4.5) becomes

$$\log A = a \log x + \log A\big|_{x=1} \qquad (4.6)$$

where for estimation purposes $\log A\big|_{x=1}$ can be treated as a constant[10]. Thus $A = cx^a$.[11]

Here the assumption of equal present value of earnings for all persons is not imposed on the model. This is an important point because it enables us to take account of the fact that entrance to professions is not completely free due, for example, to the existence of non perfect capital markets. If individuals could borrow against their future stock of human capital, then the present value of lifetime earnings would be a reasonable hypothesis. However, it has long been recognized[12] that barriers and market imperfections prevent this equalization.

The earnings equation (4.3) then becomes

$$Y_s = Y_s(H(S,A(x)),\ D,\ \varepsilon) \qquad (4.3)'$$

where an explicit form for A(x) is given by (4.6).

The set of shift parameters denoted by D contain variables such as urban/rural residence, occupational classification, self employed/employee, etc.

To use (4.3)' some assumption must be made about the way the variables are related to income. Two obvious approaches present themselves. One is to treat the relationship as multiplicative, the other is to treat it as being additive. Considering first the multiplicative case (4.3)' becomes

$$Y_s = S^{\beta_1} . cX^{\beta_2} . D^{\beta_3} . \varepsilon \tag{4.7}$$

where D^{β_3} is a vector of variables defined by

$$D' = [e^{\gamma_1 D_1}, e^{\gamma_2 D_2}, ... e^{\gamma_L D_L}] \tag{4.8}$$

The dummy variables - following normal practice - are specified in exponential form so that when a logarithmic transformation is made they enter in additive form.

To use ordinary least squares on (4.7) - after transforming by logarithms - a special assumption must be made about the disturbance term if its logarithm (ε^*) is to be distributed normally with mean zero and variance σ^2.[13] That is ε is assumed to be log normally distributed with mean $\{\exp(\sigma^2/2)\}$ and variance $\{\exp(\sigma^2).[\exp(\sigma^2)-1]\}$. To justify this, Goldberger[14] shows we can invoke the central limit theorem, which states that the distribution of the product of N independent random variables tends to log normality as N tends to infinity. Thus, if we consider ε as representing the effect of many independent multiplicative events, the assumption of log-normality is very reasonable.

Using this result, substituting (4.8) into (4.7) and taking logarithms we obtain an estimable linear form for the earnings equation

$$\text{Log } Y_s = \beta_1 \log S + c' + \beta_2 \log x + \gamma_1 D_1 + \ldots \gamma_e D_e + \varepsilon^* \quad (4.9)$$

The alternative treatment of (4.3)' is to specify an additive form for the dummies, the schooling variable and the error term ε. In this case, we obtain

$$Y_s = \alpha_1 S + cx^{\alpha_2} + \delta_1 D_1 + \ldots \delta_e D_e + \varepsilon \quad (4.10)$$

where ε is now assumed to be distributed $N(0,\sigma^2)$. Ordinary least squares is best linear unbiased for any given value of α_2 and iterations can be performed on this until a minimum sum of squares is achieved. The assumed independence of β_1, β_2 and D is examined at the estimation stage as is the choice between the functional forms (4.9) and (4.10).

4.3 Variables, Data and Estimation

Each specification makes use of the same set of variables. These are as follows: schooling, experience, occupational classification, farm/nonfarm, class of worker, area of work. None of the variables - other than income - are in continuous form and each variable has a given number of categories into which it can fall. The number and description of these categories are given in Appendix C. Thus, for example, there are eleven different occupational classifications and in the estimation these categories are treated as if they were dummy variables. So the matrix of exogenous variables will allocate eleven columns for the variable occupation and for each observation all but one of these will contain a zero entry, the remaining column having a unit entry. The micro data are in this format because of the nature of the questions asked in the survey upon which the data are based.

The "schooling", "occupation" and "area of work" variables are described in Appendix C. For estimation simplicity, the class of worker variable was aggregated into three categories: (a) employee, (b) self-employed, and (c) other. The farm/nonfarm variable defines whether or not any individual in the census family has 50% or more of his income from farm self-employment.

The inclusion of professionals, farmers etc. in the earnings equations means that the returns to education are over-estimated since some of the income is attributable to their business capital. However, since only personal assets will be considered when estimating the net asset position of the units in question, this procedure will not bias the estimates of lifetime purchasing power.

Job experience was calculated by assuming first that each educational classification corresponds to a given number of school years. This number of years plus five (i.e., the assumed school starting age) was then subtracted from the individual's age to give an approximation to the number of years for which he has been working. The approximation is due to the fact that not all individuals who fit into a given educational classification have had exactly the same number of years at school.

An alternative is to use the individual's age as a proxy for experience. If the equations were linear we should obtain the same degree of explanatory power for the equation as a whole since the experience variable is a linear combination of schooling and age. Some trial regressions with age substituted for experience yielded an almost identical R^2 but as the experience variable is preferable on theoretical grounds, these results are not used here.

The number of weeks worked has not been included as an explanatory variable in these regressions. The reason for this omission is not because it is not significant. On the contrary, it never fails to increase the explanatory power of the regressions by less than 6 percentage points. The difficulty associated with its inclusion arises from trying to predict the number of weeks an individual will work in future years. Since a prediction of this kind would involve using most of the other explanatory variables in the regression then the predicted weeks worked would be almost a linear combination of these variables and thus the inclusion of a predicted weeks worked variable could not be expected to improve the quality of the results.

4.4 Results

The results of the regressions using equations (4.9) and (4.10) are given in Tables 4.1 and 4.2. The estimation was carried out using a procedure termed "Multiple Classification Analysis" which is a subprogram of the Statistical Package for the Social Sciences designed at the Institute for Social Research at the University of Michigan. This is designed particularly for equations where the exogenous variables are in categorical form and it uses an iterative procedure rather than the conventional matrix inversion methods so as to avoid possible singularities due to the large number of dummy variables.[15]

The coefficients are estimated as deviations from the mean and thus the predicted value of the endogenous variable is the sum of the mean plus the coefficients corresponding to the categories into which an individual falls. Bearing this in mind, an analysis of the experience and education variables in Tables 4.1 and 4.2 shows that these coefficients are almost perfectly monotonically increasing in line with a priori expectations. The rate of change in these coefficients has no meaning because the difference in the value of variables between categories is not constant. The other variables also have coefficients which correspond with expectations.

The coefficients in Table 4.1 - corresponding to the additive specification - were obtained by choosing the equation which yielded the smallest sum of squared residuals while iterating over different values of the exponent entering the experience variable. The value of the coefficient thus obtained is $\alpha_2 = 1.1$.

The correlation coefficients in each case are of the order of magnitude usually obtained for this kind of cross-section work, being 33% in the additive specification and 24% in the logarithmic case. T statistics are not estimated in this program since the estimation technique does not require the inverse of the (X'X) matrix. Consequently, the diagonal elements of $\sigma^2(X'X)^{-1}$ are not available. Instead, two measures of the significance of a regressor are given. First, a one way analysis of variance is performed and the resulting F statistic is given in each case below the regressor. Second, the eta squared defines the proportion of the variance associated with the dependent variable which is explicable by the independent variable in question.

The question now arises as to which of these equations

TABLE 4.1

LINEAR SPECIFICATION

Regressor Category	Education	Potential Experience	Area of Work	Occupation	Class of Worker	Farm/ Nonfarm
(1)	-2391	-5741	758	1595	133	-68
(2)	-1597	-4928	193	410	-452	12
(3)	-508	-3044	173	-817	-1822	
(4)	219	-1544	-355	48		
(5)	684	-123	-917	-1044		
(6)	1408	781		-18		
(7)	1060	1013		-1727		
(8)	1680	857		292		
(9)	3039	493		-1148		
(10)	5723	-437		-850		
(11)				-457		
F Statistic	348.1	92.7	334.6	296.6	225.8	397.3
Eta squared	17%	5%	8%	16%	3%	3%

Dependent variable: Total earnings of male family heads.

Mean of dependent variable: 7564.1

R^2: 29.6%

Number of observations: 15379.

TABLE 4.2

LOGARITHMIC SPECIFICATION

Regressor Category	Education	Potential Experience	Area of Work	Occupation	Class of Worker	Farm/ Nonfarm
(1)	-.442	-.839	.105	.251	.145	-.028
(2)	-.225	-.542	.076	.149	-.295	.489
(3)	-.009	-.325	.05	-.002	-3.363	
(4)	.106	-.047	-.01	-.046		
(5)	.145	.117	-.167	-.112		
(6)	.212	.172		-.01		
(7)	.237	.102		-1.063		
(8)	.168	-.075		.107		
(9)	.417	-.617		-.318		
(10)	.465	-1.378		-.856		
(11)				2.405		
F Statistic	96.3	89.4	194.5	206.4	522.6	358.3
Eta squared	5%	5%	5%	12%	6%	2%

Dependent Variable: Logarithm of total earnings of male family heads.

Mean of dependent variable: 8.498

R^2: 17%

Number of observations: 15379

should be used as a predictor of future income when estimating the earnings profiles. It seems reasonable to use the equation which predicts better ex post. It does not follow, however, from this choice that the linear form should be used even though it has a higher R^2 than the logarithmic form. The logarithmic specification yields an R^2 which explains a certain percentage of the variation of the log of income from its mean. Thus, it does not explain the same variation as the additive equation. A retransformation was thus applied to the logarithmic form (i.e., the antilog of the predicted value of log Y was obtained) to obtain an estimated value for income. A correlation between this prediction and the actual value of income yielded a smaller R^2. This confirmed that the additive specification is more appropriate than the logarithmic for prediction purposes.[16]

In leaving the specification issue, it is worth pointing out that the estimation of (4.9) is equivalent to regressing log Y_s on the regressors specified in equation (4.10). This is because all right hand side variables must be in categorical form. This is the actual estimation procedure which has been used.

4.4.1 Collinearity and Specification

It was pointed out earlier that years of schooling may be correlated with experience. This may mean that the coefficient estimates are not precise. However, in this instance the degrees of freedom are considerable and thus the usual problems associated with using ordinary least squares in the presence of collinearity are unlikely to arise.[17]

A more serious possibility, as suggested in footnote 10, is that equation (4.10) may be misspecified because no slope dummies have been included to allow for the possible dependence between education and experience. This possibility was tested for in the following way.

Equation (4.10) was estimated (a) allowing for this interaction and (b) ignoring it. If R_a^2 and R_b^2 represent the proportion of the variation explained in each case then a test of significance of the extra k variables can be performed by making use of the fact that $\frac{(R_a^2 - R_b^2)/k}{(1 - R_a^2)/(n-k)}$ is distributed as an F statistic with (k, n-k) degrees of freedom, where n is the number of observations. This statistic was evaluated at 1.45 making it significant at the 5% level but not at the 1% level. As the question of including or excluding the interaction was a marginal one, it was decided to exclude it on grounds of simplicity as the computation of future incomes would be much more complex in a situation where there are about 110 coefficients than in a situation where there are 'only' 40 coefficients.

4.4.2 The Residuals

Given the wide range of values observed in the sample it would be unusual to find homoskedastic residuals. If the residuals are heteroskedastic then ordinary least squares is not efficient. To test for this, the Goldfeld-Quandt test was used.

The data were divided into two subsamples. The first contained 7649 observations and was restricted to values of earnings less than $7,100. The second subsample contained 6054 observations and was restricted to values of the dependent variable greater than $8,100. Thus, this involved omitting 1676 observations in the middle of the data following Goldfeld and Quandt's suggestion that this procedure may increase the power of the test.

The statistic $e_a'e_a/e_b'e_b$ was estimated - the e_i denoting residuals - and its value was 4.37. This is distributed as an F statistic with (n_a-k, n_b-k) degrees of freedom (k is the number of regressors and n_i the number of observations in each regression). The critical values on the F table corresponding to this are approximately 1.0 at both the 1% and 5% significance levels. Consequently, the hypothesis of homoskedastic error terms could not be maintained and the parameter estimates in Table 4.1 are thus not efficient - but are unbiased and consistent.

The usual procedure for obtaining efficient unbiased estimates is to use Generalized Least-Squares — which is equivalent in most cases to transforming the data by some function of an independent variable with which the error term is correlated. This approach is not possible here since the regressors are not in continuous form. A large sample approximately efficient method is given by Theil.[18] However, this was not pursued since with about fifteen thousand degrees of freedom, the consistency properties of ordinary least squares can be called upon.

To conclude: The objective of this chapter was to estimate earnings equations for males so that life-cycle earnings could be predicted. The coefficients to be used for this purpose are given in Table 4.1.

NOTES ON CHAPTER FOUR

1. For a description of the date see Appendix A.

2. See for example Becker (1964), Mincer (1974) and Hanoch (1967).

3. For example Taubmann and Wales (1974).

4. Mincer (1974), Eckaus, El Safty and Norman (1974).

5. Stiglitz (1975) provides a good review of this literature.

6. If income is specified to grow at an annual rate g the analysis still carries through with the result that $Y_s/Y_{s-d} = \exp(r-g)d$.

7. The results of this are given in Appendix D.

8. Becker (1964).

9. For evidence on this see Hanoch (1967).

10. In the foregoing the rate of change of A has not been specified to depend on the number of years of schooling. This modification will be incorportated at the estimation stage by interacting the variables A and S.

11. It should be pointed out at this stage that we do not distinguish between firm specific and job specific training/experience. This is because of the manner in which the life cycle earnings are estimated i.e. we do not have information on the individuals' jobs or firms for which they work, only on the general nature of their profession (see Appendix A).

12. Friedman and Kuznets (1954) were among the first to estimate the differentials in earnings which could be attributed to barriers to entry. They calculated that as much as 50% of the difference between the earnings of dentists and physicians is attributable to the greater difficulty associated with entering the medical profession (p. 395).

13. The normality assumption insures that O.L.S. is also maximum likelihood.

14. Goldberger, "Econometric Theory", Wiley 1964, p. 216.

15. Further details can be found in Andrews et al (1973).

16. Note that the distribution of the ration of the coefficients is not known and thus we cannot say if one fit is significantly better than the other. This fortunately does not prevent a valid choice being made on the basis of the best predictor criterion.

17. For a discussion of this see Johnson (1972) pages 159-170.

18. See Theil (1971) p. 245 or Kmenta (1971) p. 261 for a detailed description.

CHAPTER 5

LIFE-CYCLE EARNINGS FOR FEMALES

In this chapter, we estimate life-cycle earnings for females. The chapter proceeds as follows: In the first section, we outline a method of deciding upon whether the woman works or not. In the second section, we estimate how much she will earn if she is predicted to work. In the final section, we examine the results.

5.1 The Participation Decision

When predicting the future income stream for the females there are several approaches which could be adopted. One is to estimate expected income in one stage - using in the sample both zero and non-zero income earners. Another is to use a multistage procedure in which participation in the labour force, or weeks worked, is first determined and on the basis of this to estimate an annual or weekly earnings figure. The latter procedure is more reasonable given that about 60% of married women do not participate in the labour force and it is the approach we adopt.

Following this line of argument, we estimate a system of equations of the following type.

$$LPF = LPF\ (X_1, X_2, \ldots\ldots X_k) + \varepsilon \qquad (5.1)$$

$$Y_w = Y_w\ (X_{k+1}, X_{k+2}, \ldots\ldots X_m) + \varepsilon \qquad (5.2)$$

where LPF denotes labour force participation and Y_w the female's income. The X vectors define sets of variables which determine participation and income respectively. Effectively then, we are estimating a recursive model of family labour supply. The head, if male, is assumed to work full time until retirement while his wife's decision depends on the variables defined in (5.1), among which is the head's income.

The estimation of (5.1) could be performed using ordinary least squares where the dependent variable takes on the limited values of zero or one. The treatment of such procedures has been investigated by Goldberger.[1] He shows that a major drawback is that the assumption of homoskedasticity is untenable and thus generalized least squares is the appropriate technique. However, even if this can be successfully applied, there remains the difficulty that predicted values can fall outside the interval between zero and one and this can be shown to be inconsistent with the assumptions required for the use of generalized least squares.

A more attractive procedure is to use discriminant analysis as it does not suffer from the shortcomings of ordinary or generalized least squares. It is designed for the following type of problem.

Assume we have k populations and samples drawn from each. Discriminant analysis involves formulating a rule - constructed by using the information contained in the samples - which enables us to assign a new observation to one of the k populations. The procedure involves minimizing a loss function.[2]

Thus, in the labour force participation problem, we observe drawings from two populations - those who work and those who do not. Each drawing gives information on the characteristics of one individual, e.g. age, education, number of children, etc. We can look upon these characteristics as the set of variables in equation (5.1) above.

Formally, let us denote the two populations by P_1 and P_2 and the vector of characteristics by $X = [x_1 x_2 \ldots x_k]$. Now define two density functions $f_1(X)$ and $f_2(X)$ which characterize the distribution of the vector of personal characteristics X and define p_1 and p_2 as the a priori probabilities that a drawing comes from population one or two respectively.

Now divide the space X into two regions S_1 and S_2 in such a way that if X falls in S_1, we assign the individual to P_1 and if X falls in S_2 we assign to P_2. Finally, we define C as the cost of misclassifying an individual. The objective of the analysis is to minimize the expected costs of misclassification.

The probability that an individual from P_1 is wrongly classified is given by

$$\int_{S_2} f_1(X)\ dX \tag{5.3}$$

Thus, the overall expected costs of misclassification are given by

$$C\ p_1 \int_{S_2} f_1(X)\,dX + C\ p_2 \int_{S_1} f_2(X)\,dX \tag{5.4}$$

Using the fact that

$$\int_{S_1} f_2(X)\,dX = \int f_2(X)\,dX - \int_{S_2} f_2(X)\,dX \tag{5.5}$$

the expected costs can be written as

$$C \int_{S_2} [p_1 f_1(X) - p_2 f_2(X)]\,dx + C\ p_2 \int f_2(X)\,dX \tag{5.6}$$

Since the last term is a positive constant and independent of S_2, then equation (5.6) is minimized if S_2 is defined as the set of X's for which $[p_1 f_1(X) - p_2 f_2(X)]$ is less than zero, and thus S_1 as the set of X's which makes this positive. Thus, we assign a drawing to P_2 if

$$f_1(X)/f_2(X) \leq p_2/p_1 \tag{5.7}$$

otherwise we assign to P_1.

At this stage, we must make some assumption regarding the density functions $f_1(X)$ and $f_2(X)$. The most common procedure (e.g. Adelman and Morris (1968)) is to assume each is multivariate normal with the same covariance matrix Σ but different mean vectors μ_1 and μ_2. The ratio of the densities, after simplifying, becomes

$$\frac{f_1(X)}{f_2(X)} = \exp\{X'\Sigma^{-1}(\mu_1-\mu_2) - \tfrac{1}{2}(\mu_1+\mu_2)'\ \Sigma^{-1}(\mu_1-\mu_2)\} \qquad (5.8)$$

Rewriting $\Sigma^{-1}(\mu_1-\mu_2) = \theta$ and taking logs of both sides the rule becomes: assign to P_2 if

$$X'\theta - \tfrac{1}{2}(\mu_1+\mu_2)'\ \theta \le \log(p_2/p_1) \qquad (5.9)$$

otherwise assign to P_1.

Maximum likelihood estimates of μ_i, p_i and Σ can be obtained from the sample observations. μ_1 and μ_2 are estimated by using the sample means $\bar{X}_1$ and $\bar{X}_2$. The maximum likelihood estimate of Σ is given by

$$S = [1/(n_1+n_2-2)][X_1'X_1 + X_2'X_2]$$

where n_1 and n_2 are the number of observations in each sample and X is now a matrix of <u>deviations</u> from the appropriate sample means.

So for each individual we estimate $X'\theta$ (termed the discriminant score) by substituting the observed values for X. This returns a scalar and depending on whether it is greater or less than $[\log(p_2/p_1) + 1/2\ (\mu_1+\mu_2)'\ \theta]$ the individual is assigned to be a participant or a non-participant in the labour force.

5.1.1 Data

The data are taken from the 1972 survey.[3] A sample of three thousand was taken from the tape and of these about two thirds were families in which both parents were alive.

Of these 61% of the females did not participate and 39% did. These rates were thus used as estimates of p_1 and p_2 in equation (5.9) above.

Participants were considered to be any females who worked for at least one week in the year. Most, however, worked full time both in terms of weeks per year and hours per week. The figure for the former is 68% and for the latter 61%.

Difficulties associated with defining participation arise in dealing with expected income. By including part time workers we will probably overestimate their income and consequently underestimate the income of full time workers. The alternative is to estimate an intermediate equation to determine the number of weeks per year the female will work. But given that such a procedure would involve considerable complexity for a very low return (due to the nature of the problems associated with such which we have already discussed in Chapter 4) this is not a practical extension of the two equation model.

In computing the discriminant score, the following set of variables was used:

1) education level;
2) number of dependent children;
3) number of children under five years of age;
4) urban/rural residence;
5) income/earnings of husband;
6) age;
7) marital status.

The coefficient estimates (i.e. the elements of θ are given in Table 5.1.)

TABLE 5.1

Variable	Coefficient Value
Education	.0042
Number of children	-.1084
Number..under five	-.5996
Residence	-.0866
Marital Status	6.3311
Age	-.1292
Education of head	-.0154
Age of head	.0922

It should be pointed out that the interpretation of these coefficients is not the same as in regression analysis to the extent that the value of $X'\theta$ must be compared with the right hand side of equation (5.9) and this was estimated to be 4.55. A smaller coefficient does not mean that the variable is of lesser significance than one with a larger coefficient since the critical magnitude is the product of the coefficient and the variable.

The variables 'education', 'number of children' and 'age' require no explanation as their direction of influence is well recognized. The residence variable takes on a value of (a) zero, or (b) one depending on whether the family lives (a) on a farm, or (b) other respectively.

The woman's occupation was not used as a variable in determining participation. This variable should influence the work/leisure decision both because it might serve as a proxy for income and because it would be a measure of the attractiveness or otherwise of available work opportunities. The reason for not using it is that the data yielded no occupational classification for the vast majority of those who did not participate in the labour force. Of the nonparticipants sampled, over 90% did not state an occupation, while the corresponding number for participants was about 27%. Thus there was no way of incorporating this information into the model even though it would be desirable to do so.

The marital status variable is included to allow for differential work rates between married and widowed females. In practice, this difference is considerable.

The specification of the income or earnings of the male is not a minor problem. We could use the husband's earnings for the year in which the female's decision is being made. Alternatively, his total income could be used or even his lifetime earnings or income. Consider the choice between current earnings and income.

On theoretical grounds, a case can be made for using either. If total income is known with certainty then this is the appropriate measure. But if nonlabour income is not known with the same degree of certainty as labour income then earnings should be a better predictor of participation rates than income. Both variables were used here in separate estimations and there was no difference in their ability to predict ex post. Whether this is due to the fact that some women make their decision on the basis of the head's income and some on his earnings with the resulting aggregation being indecisive or simply that there is a strong correlation between earned and unearned income it is not possible to determine.

Since an earnings profile must be estimated for the household head to obtain a measure of lifetime wealth while a profile of lifetime income is not necessary, it was decided to use earnings.[4]

So far we have taken no account of the wage or earnings which the female could expect to earn if she were to participate in the labour force. Initially, this may seem a serious omission given the theory of participation which has evolved based on the accumulation of human capital.[5]

However, we do not have an observed wage offer for all possible participants. This would thus have to be estimated from the data available for those who do work. The bias associated with this procedure is that it assumes those who don't work face the same potential wage rate as those who do - for a given set of characteristics. But the average wage of those who do participate in the labour force may be an upward biased measure of the average wage offer for all members of a group. Consequently, imputing a wage to nonworking women estimated from a sample of those who do work would introduce a "selectivity bias".[6]

Even if this bias could somehow be overcome, it would still not increase the predictive power of the discriminant analysis. This is because essentially the same set of variables determine both income and participation. So, if a wage or income offer were to be estimated, it would be done using approximately the same set of variables used in the discriminant analysis. But this estimate is then equivalent to taking some linear combination of the variables already being utilized and using this alongside those variables. It is clear that this would not improve the predictive power of the analysis and since it is prediction rather than labour supply elasticities we are concerned with, there is no reason to include an imputed income or wage. This is the same issue encountered in using a "weeks worked" variable to explain male earnings. It seems that the only alternative way of estimating a wage would be to utilize data on occupational classification, but the difficulties associated with this have already been alluded to.

5.2 Earnings Equations

The estimation of earnings equations for females poses considerably greater problems than for males. The difficulties arise first because of the nature of the data available and second because there seems to be a greater random component associated with annual female earnings unless we can accurately predict whether they will work full or part time - both in terms of hours per week and weeks per year.

The first problem arises because of the fact that occupational data are not available for women who don't work. Thus, even in cases where individuals not currently employed are predicted to be employed in the future, the absence of data on occupational classification will mean that the predictive power of the equations will not be as strong as otherwise.

The second issue alluded to above is probably the more serious of the two - that is how to treat part time workers. The inclusion of weeks worked per year and hours worked per week as explanatory variables will improve the fit of the equation when predicting annual earnings. But while there is no difficulty in obtaining base year data for these variables, a significant problem arises when it comes to making assumptions regarding part time activity in future years.

The approach adopted here is to omit this variable. While this initially may seem to be a serious omission, it can be argued that its shortcomings are less than those associated with the alternatives. Let us consider these.

One possibility is that alluded to earlier; i.e. to predict whether a given individual will be working part time or full time in any given year and then include this prediction as the appropriate variable. As before however, we face the same limitations on data availability.

A second solution would be to assume that the individuals in question will continue to work the same number of hours in all future years as they did in the base year. But, given changes in family structure over time, this would not be a very fruitful approach either.

Another solution is to omit the part time variables at the estimation stage and then to allow for their effect at the profile estimation stage by assuming that some fraction of the difference between predicted and actual earnings can be attributed to the fact that many individuals work part time. The difficulty with this is that the estimated regression coefficients will be biased because of the specification error unless the omitted variables happen to be orthogonal to all included variables[7], and this is extremely unlikely.

Thus, due to data limitations and econometric problems, there seems to be no ideal way to estimate future earnings which explicitly takes account of the full time/part time work decision. However, the omission of this variable does not bias the estimates of lifetime family earnings because the correlation between male and female earnings is extremely weak.[8]

In addition to the above, there is an additional issue concerning the experience variable which was not encountered in estimating the earnings equations for males. In the male earnings equations, it was assumed that experience could be approximated by the difference between the individual's age and the estimated age at which he joined the labour force.

For females, however, this is not a valid procedure because of their more frequent entry/exit to and from the labour force. It is thus not appropriate to assume that two women of the same age have equal work experience if they graduated at the same age whereas this is a valid assumption for males.

In view of this, we use the number of children in the family as a proxy for the number of spells during which the woman was not working. This, in conjunction with age, serves the same purpose as the experience variable in the earnings equations for males.

A marital status variable has been included in that it partly overcomes the full time/part time problem since we would not expect the work pattern of those married to be the same as that of those who are single, divorced or widowed.

As in the case of the male earnings equations, there seems to be no a priori reason as to why the variables should determine earnings in multiplicative or additive form. Thus, both linear and loglinear equations have been estimated and the results for each are given in Tables 5.2 and 5.3 respectively.

5.3 Results

In terms of explanatory power, the results are good; the explained portion of the sum of squares being 35% and 41% in the linear and loglinear specifications respectively. The inclusion of weeks worked as an explanatory variable increased the predictive power of the equations by about 5% in each case. It is thus unfortunate that this could not be incorporated into the earnings profiles in a satisfactory manner.

Considering Tables 5.2 and 5.3, we see that the 'education', 'number of children' and 'age' variables exhibit almost perfectly monotonic patterns. The negative signs associated with the high age brackets are probably caused by the phenomenon discussed in Chapter 4, section 4.1. The 'occupation', 'area of work' and 'class of worker' variables are not ordered in such a way that we would expect particular sign patterns, but given the categories to which they correspond (see Appendix C) their sign and relative magnitude correspond to what we would expect a priori.

The correlation coefficient is higher for the loglinear specification than the linear specification and the question thus arises as to which of the equations should be utilized when estimating the earnings profiles. The correlation between the antilog of the predicted values of earnings resulting from the logarithmic equation and actual earnings was estimated and found to be considerably less than the R^2 for the linear equation. On the basis of ex post predictive ability, the linear equation was thus considered to be more appropriate

TABLE 5.2

Regressor	Education	Occupation	Area of Work	Class of Worker	Number of Children	Age	Marital Status
Category							
0					477		
1	-745	573	310	722	57	-904	-55
2	-469	1335	-30	-550	-441	-215	417
3	-230	-89	-267	-1764	-662	118	
4	143	-1308	-26	-1907	-863	163	
5	62	-1118	-398	-1703	-747	300	
6	140	-165			-1281	272	
7	421	-594			-996	89	
8	156	-353			-1221	36	
9	838	-113			-774*	-221	
10	1992	-1648			1306*	-631	
11		429			-148*		
F Statistic	121	240	90	366	28	8	21
Eta Squared	14%	26%	5%	16%	4%	1%	1%

Dependent Variable: Total Earnings of Females

R^2: 35%

Number of Observations: 7609

Mean of Dependent Variable: 3081.7

*The number of observations in these categories is less than 4 in each case

TABLE 5.3

Regressor	Education	Occupation	Area of Work	Class of Worker	Number of Children	Age	Marital Status
Category							
0					.248		
1	-.794	.261	.198	.850	.117	-.028	-.03
2	-.369	.235	.215	-1.672	-.246	-.084	.221
3	-.147	-.002	-.077	-5.108	-.362	-.056	
4	.241	-.492	.106	-.814	-.5	-.09	
5	.131	-.627	-.455	-1.26	-.554	.221	
6	.278	-.042			-1.254	.088	
7	.398	-.469			-.949	-.085	
8	.162	.261			-1.159	-.144	
9	.596	.285			-3.597*	-.244	
10	.63	-1.382			-1.522*	-.615	
11		.303			1.691*		
F Statistic	49	204	158	1000	20	11	14
Eta Squared	6%	23%	8%	35%	3%	2%	1%

Dependent Variable: Logarithm of total earnings of females.

R^2: 41%

Number of Observations: 7609

Mean of Dependent Variable: 6.831

*The number of observations in these categories is less than 4 in each case for the profiles.

To conclude: Our objective in this chapter was to build a model which would enable us to predict life-cycle earnings for females. We have constructed a two stage model in which we first predict whether or not an individual will work. This is done using discriminant analysis. The second stage is to predict an earnings figure for those who work. The coefficients used for this purpose are those in Table 5.2.

NOTES ON CHAPTER FIVE

1. See Goldberger (1964) p. 248-251.

2. See Dhrymes (1970) p. 65 or Johnston (1972) p. 334 for a more detailed discussion.

3. See appendix C for a description of these data.

4. At this state a computing factor dictated the course of things. Due to the large core requirements of the programmes the estimation of the male and female had to be done seperately and there was thus no easy way to use the lifetime male earnings as a determinant of the female income. Instead some of the factors determining male earnings were used and this is why the age and education of the household head appear in table (5.1).

5. Becker (1964), Mincer (1972, 1974) or Michael (1972).

6. For a more detailed account of these difficulties see DaVanzo, DeTray and Greenberg (1973).

7. See Theil (1971) p. 549.

CHAPTER 6

THE ANNUITY DISTRIBUTION

In this chapter, we use the results of Chapters 4 and 5 to derive the annuity distribution. The chapter follows this pattern: We start by examining a technical problem related to the degree of variation in life-cycle earnings. We then describe the programming of the annuities. In the third section, we show how to incorporate family size explicitly into the distribution. Finally, we examine the estimation of net worth of the families in the sample.

6.1 Optimal Variation

Since the objective of this study is to examine the variation in lifetime wealth across households, it is necessary to insure that in the process of estimating lifetime earnings the variation which actually exists is not artificially reduced. A major statistical problem exists when predicting earnings for households in any given year using the coefficients estimated in the regression equations. This arises from the fact that the regression coefficients yield point estimates of earnings. The direct use of these values means that if we have several households with the same set of socio-economic characteristics, each of these households will be assigned an identical earnings estimate. In practice, though, this is never the case. An analysis of the cross-section indicates that considerable variation exists among households which are for example of the same vintage, where the head has the same occupation etc..

If we were simply interested in the expected earnings of an individual or family at some time in the future, then the results of the regression equations would predict this - subject to the usual assumptions of the classical linear model. But since our interest lies in the variation among household incomes, it is clearly invalid to assign an identical income to all households which have the same set of characteristics.

The result which such a procedure would have on the measured (in)equality can easily be seen by examining some summary index of inequality - for example the Gini index. This is defined by

$$G = (1/(2n^2m))\Sigma_{i=1}^{n} \Sigma_{j=1}^{n} |Y_i - Y_j| \qquad (6.1)$$

where m is the mean, n the number of households and Y_k the income of household k. By assigning the same estimated income to like households, this would decrease the number of absolute differences in the Y_k (as many of these terms now become zero) and thus decrease the value of the index. This would give the impression that a lower level of inequality exists than is the case because of this statistical anomaly.

To overcome this difficulty, we must disperse the predicted values in each year about the point estimates and do this in such a way that the variance associated with the incomes of a like group of households corresponds with that of the sample.

Consider now a group of households in a given socio-economic group. Here, we will define a socio-economic group as one in which the variance associated with the random component of earnings is constant. Formally, a predicted value of (future) earnings, $\hat{Y}_i$, is estimated for each person in the group. This estimate is based on the underlying model

$$Y = XB+\varepsilon \qquad (6.2)$$

where X is a matrix of exogenous variables, B is a vector of coefficients and ε a random component where each element is identically distributed with zero mean and constant variance σ^2.

Our objective here is to derive conditions governing the size of random elements such that when the predicted values of earnings are adjusted by these terms, the variation associated with these adjusted predictions will be the same as the variation in the actual data. We will approach this problem in two stages. First, we will derive these conditions for the base year (in this case the X matrix determining the <u>predicted</u> value of earnings is exactly the same as the matrix determining the actual values). Then we will derive the conditions for the future years (here the equality of the X matrices does not hold exactly).

The variance in the actual data in the base year in any given cell (this defines a socio-economic group) conditional on the values of $\bar{Y}$ (a vector containing the mean of Y) is given by

$$
\begin{aligned}
& 1/N\ [(Y-\bar{Y})'(Y-\bar{Y})] \\
= \quad & 1/N\ Y'Y-\bar{Y}^2 \qquad (6.3)
\end{aligned}
$$

Substituting for Y using equation (6.2) above, this can be shown to reduce to

$$1/N\ \hat{B}'X'X\hat{B} + 1/N\ \hat{\varepsilon}'\hat{\varepsilon} - \bar{Y}^2 \qquad (6.4)$$

where $\hat{B}$ is the ordinary least squares estimate of B and $\hat{\varepsilon}$ is a vector of residuals.

The variance in the predicted values (including the random component η), conditional on $\hat{Y}$ and $\bar{Y}$, is given by

$$1/N\ [(\hat{Y} + \eta-\bar{Y})'(\hat{Y} + \eta-\bar{Y})] \qquad (6.5)$$

Expanding (6.5), by using the relation $\hat{Y} = X\hat{B}$, taking its expected value[1], and using the fact that the elements of η are independent of $\hat{Y}$ and $\bar{Y}$ we obtain

$$E\ 1/N\ \hat{Y}'\hat{Y} + E\ 1/N\ \eta'\eta + E\ 1/N\ \overline{Y}'\overline{Y} - E\ 2/N\ \hat{Y}'\overline{Y}$$

$$= 1/N\ \hat{B}'X'X\hat{B} + E\ 1/N\ \eta'\eta - \overline{Y}^2 \qquad (6.6)$$

To insure the equality of (6.4) and (6.6), it is clear that we need

$$E\ 1/N\ \eta'\eta = 1/N\ \hat{\varepsilon}'\hat{\varepsilon}$$
$$= 1/N\ (N-k)\ S^2 \qquad (6.7)$$

where k is the number of regressors.

Now define the variance of each element of η by v. It follows that the random numbers must be generated such that the variance equals $(\hat{\varepsilon}'\hat{\varepsilon}/N)$.

Before going on to the second stage of this problem, one further point should be noted.

It could be argued that the conditions on the generation of the random numbers should not assume $\overline{Y}$ and $\hat{Y}$ fixed. It is shown in Appendix F that the relaxation of this assumption does not change the foregoing conclusion, defined in equation (6.7).

The second stage of the problem can best be introduced by considering what we mean by a given socio-economic group. We have already defined this as one in which the variance associated with all incomes is the same. In addition to this requirement, we also define such a group as one which encompasses people of the same age and occupation. It is important to note that this does not mean that all families in a group are identical in the sense that every characteristic which defines them is identical. For example, people with the same age and occupation can have self employed or employee status and still be in the same group. Ideally, we would like to have a multidimensional classification such that only people who are identical in every respect would fit in a given cell. The reason we do not follow this procedure is as follows. Inspection of Table 4.1 indicates that there are approximately 33,000 possible combinations of characteristics. But as there are only 25,000 observations, it is clear that there is no possibility of finding people with all possible combinations of characteristics let alone have enough degrees of freedom to estimate a variance for each of these cells. The implication of these data limitations is that families must be grouped into a limited number of categories. The procedure followed here is to classify families by age and occupation of the head, where there are 10 age and 11 occupation groups. This is illustrated in Figure 6.1.

Thus, according to the procedure outlined above, we estimate a value of earnings for each individual in each cell and add to this the random element which has been generated subject to the required conditions. As people age, they move from one cell to another and their earnings are subject to a

FIGURE 6.1

Age		1	2	3	 10
Occupation	1				
	2				
	3		.		
	11				

different variance - usually larger. At this stage, the problem caused by data limitations arises.

When households move, for example, from cell (1,1) to cell (1,2), their predicted earnings change. As well as this, we know that the variance associated with the earnings of households in cell (1,2) is not the same as the variance of the earnings of the other group in cell (1,1). Thus, when the group in cell (1,1) ages by one unit, we assume that the variance associated with their new predicted earnings is the same as the variance associated with the earnings of the group who have just left cell (1,2) and progressed to cell (1,3). This latter group likewise will have their predicted earnings subjected to a new variance. The rationale for their progression is that we believe people with different earnings levels will likewise have differing degrees of variation in their earnings.

This procedure has some merit to the extent that it generates results which correspond to observed patterns. There is, however, a problem associated with it and this arises because of the fact that not all individuals in a given cell are identical in every respect. For example, the actual earnings variation in cell (1,2) is attributable in part to the fact that some individuals are self employed while others are employees, some live in urban areas while others live in rural areas, some have university education while others have high school and so forth. Thus, to the extent that some of the variation is attributable to these differences coupled with the fact that the group coming into this cell will not have exactly the same set of demographic and economic characteristics as the group for which the variance was estimated (though they do have the same age and occupation), this will mean that the conditions under which the random numbers were generated will not be exact. This approximation, however, is unavoidable due to the data limitations alluded to earlier.

Finally in this section, we consider the exact specification of the generation of the random numbers. The numbers used are initially distributed uniformly on the interval (0,1). Denote these numbers by γ_i. In order that they have zero mean and lie on the interval (-1,+1) they are transformed by the factor $(2\gamma - 1)$. The variance of this set of numbers is

$$\int_{-1}^{+1} (x-m)\ f(x)\ dx \tag{6.8}$$

where x is $(2\gamma - 1)$, m is the mean of x and f(x) the density function defined by 1/(a-b) where a and b are the limits of integration. It is easily shown that (6.8) has a value of 1/3. In order that the numbers have unit variance, we thus scale them by $\sqrt{3}$. We can define $\theta = \sqrt{3}(2\gamma-1)$ as a uniformly distributed random variable with zero mean and unit variance.

If the numbers are to have a variance defined by equation (6.7) i.e. $\hat{\varepsilon}'\hat{\varepsilon}/N$ it follows that we merely multiply each θ_i by $\sqrt{\hat{\varepsilon}'\hat{\varepsilon}/N}$ yielding the required result[2]

$$\eta_i = \sqrt{3}(2\gamma_i-1)\ \sqrt{\hat{\varepsilon}'\hat{\varepsilon}/N} \tag{6.9}$$

6.2 Programming of Earnings Profiles

6.2.1 Male Earnings

An earnings figure is estimated for each member of the sample for each year and discounted to the base year. This predicted value includes the random component described in the previous section. In doing this, an allowance is also made for secular growth in the economy by boosting the predicted value of earnings by a factor of 1.02 for each year.[3]

The random component is also boosted by the same factor with the result that the variation in earnings for a given individual will increase through time by more than would be the case if he moved from one age bracket to another in the cross-section.

To determine whether any person will live for another year, a uniformly distributed random number is generated and compared with the known probability of living at each age. Since the probability of death increases each year after a certain age, all of the members in the sample die off in time.

All persons less than sixty five years of age who are working in the base year are assumed to work full-time until that age and then retire. Those who are already retired in the base year or are classified to be permanently unable to work are assigned an income equal to that which they receive in the base year (boosted again at the rate of two percent per annum) for every year until they die. Those working between the ages of 65 and 70 in the base year are assumed to retire at the age of 70. Those working over the age of 70 are assumed to retire the following year.

The question of retirement income and pension plan contributions is a difficult one in the absence of data for this. There are two possible courses of action here. One is to neglect contributions during the working life and thus assume that the present value of pension income will be exactly offset by not deducting this amount. The difficulty with this is that since we do not capture each person in the sample at the beginning of their working life, it would mean that we would be neglecting future income which accrues because of past deductions from earnings. Given this, the best alternative is to deduct a certain percentage from earnings in each year and assign a certain income to each person on retirement. The procedure which has been adopted here is to deduct 5% of gross earnings from all employees and 10% from those who are self-employed each year and on the age of retirement to assign an income equal to one half of the amount they earned in their final year of work.

In each year, there is a non-zero probability that any individual will be unemployed. There are again various ways to allow for the effect of this on predicted earnings. One is to use in the earnings equations only those persons who are employed full time and subsequently to establish a probability and duration of unemployment and adjust the predicted earnings accordingly. This effectively yields an expected value of earnings. A simpler procedure which yields the same result - in terms of expected earnings - is to include in the earnings equations those who are not employed for the whole year and the predicted values in this case incorporate the effect of the possibility of being unemployed. It is this procedure which has been adopted here.

6.2.2 Female Profiles

The same general approach which has been outlined above has also been adopted for predicting female earnings. Each year, the probability of death is known and this is compared with a random number drawn from the uniform distribution. If the female lives, we then estimate the discriminant function to decide on whether she will work or not.

Between these two stages the family structure is examined since the size of the family or any increment to it will affect the work decision. To simulate family size, we compare the actual with the average family size and if the number of dependents is less than the average, the family size is increased by one. Each year, the same process is used but the family is never augmented beyond two. Constraints are also built into the program to permit increases in family size only (a) when both parents are alive and (b) when the female is less than 50 years of age.

Given the family size, this enables a discriminant score to be estimated. If this exceeds the critical value, the woman is assumed to work, if it is less than the critical value she is assumed not to work. In the latter case, the assigned income is zero. If she is predicted to work, she is assigned an income which is determined in the same way as that for the male - i.e. a predicted value is obtained from the earnings equation and a random component is added to disperse these values.

Children's allowances have been included as part of the total income of the female. An ordinary least squares equation was estimated, regressing payments on the number of members in the family. The data were taken from The Survey of Consumer Finances (1972). (See Appendix G.)

6.2.3 Imperfect Capital Markets

It has long been recognized that capital markets are not perfect despite the frequent use of such an assumption. In practice, the lending rate differs from the borrowing rate and the rate at which lenders lend to borrowers depends on the capital stock (both human and nonhuman) of the borrower.

The life-cycle model which has been developed in the previous sections has not taken account of this explicitly. A recent paper by Applebaum and Harris (1978) makes use of the fact that capital markets are not perfect in an attempt to reconcile the observed pattern of household consumption with the implications of optimal life-cycle behaviour, as developed by Strotz (1956), Yaari (1964) and others. The theoretical work implies that consumption should be monotonically increasing or decreasing over the life-cycle while observed patterns are 'hump' shaped with the peak occurring in the middle years. Applebaum and Harris show that if the interest rate depends on accumulated capital, then it is possible to reconcile the theory with fact. In terms of the model developed in Chapter 3, equation (3.9a) becomes

$$\frac{\dot{C}}{C} = \frac{r(K) + K\frac{\partial r}{\partial K} - \delta}{(1-\beta)}$$

Clearly $\dot{C}/C$ is now not necessarily monotonic over the life-cycle.

The omission of this factor from the estimation of life-time wealth would mean that the variation in the distribution would be underestimated. This is because those families at the lower end of the distribution face a higher interest rate when borrowing than those at the top because the former have a lower stock of capital. It is desirable to capture this effect in the annuity estimation. It is important to point out however, that of the two differentials referred to above, i.e. differential between borrowing and lending rates and borrowing rate depending on accumulated capital, the second is likely to be more important. The reason is that the major purpose for borrowing in a lifetime is the purchase of a house. But since the house itself is the collateral in such cases, the interest rate charged to a low income household will differ little from that charged to a high income household. The same argument applies in many cases for automobiles and durables.

Consequently, it can be argued that, even though the Applebaum-Harris model can generate a hump-shaped consumption pattern, the practical implications are unlikely to be significant.

Despite this there is still a strong case for using different discount rates for low and high income families. This is because incomes are positively correlated with age and debt negatively correlated with age. Thus, when major debts are paid off, the effective discount rate changes. Instead of paying a rate of 11% on mortgages, the family finds it can earn perhaps 7% on savings.

To take account of both differentials outlined above, the interest used in the estimation of the annuities decreases with the age of the family head by a fraction of one percentage point each year until the age of fifty. In technical terms, this simply means that in formulae (6.23) and (6.24) the discount rate is no longer a constant from one time period to the next.

6.3 Family Size

In Chapter 3, it was shown that an adequate treatment of the annuity estimation requires a consideration of family structure. At that stage, it was shown that family size could be incorporated into the analysis explicitly by re-writing the utility functional as

$$V = \int_0^T F_t U(C_t/F_t) e^{-\delta t}\, dt \tag{6.10}$$

or

$$V = \int_0^T U(C_t/F_t) e^{-\delta t}\, dt \tag{6.11}$$

Before considering how the specification of the annuities changes when (6.10) or (6.11) is used instead of (3.1), we deal with a general problem which arises when we have two variables (consumption and family size) in the utility function. Consider (6.10).

The instantaneous constraint governing the maximization of (6.10) is that consumption (C) plus savings (S) must equal income (Y) plus the return on capital (K). The stock of capital (K) and family size (F) are the state variables. Savings (S) and $\dot{F}$ are the control variables, where $\dot{F}$ is the time derivative of family size.[4] The Hamiltonian is then

$$H = e^{-\delta t} FU \left(\frac{C}{F}\right) + \lambda_1 S + \lambda_2 \dot{F}$$

where λ_1 and λ_2 are the costate variables and time subscripts have been omitted for convenience. The conditions

$$\partial H/\partial S = 0 \quad : \quad \partial H/\partial \dot{F} = 0$$

$$\partial H/\partial K = -\dot{\lambda}_1 \quad : \quad \partial H/\partial F = -\dot{\lambda}_2$$

define the necessary conditions for an interior maximum.

Our interest is in estimating numerical values for annuities, not in explaining family size, and since the solution to a two state, two control problem becomes very unwieldy we will at this stage assume that the optimization with respect to family size has already been carried out and that the values in (6.10) represent optimal values. Equation (6.10) can then be viewed as a 'concentrated' utility functional and this enables us to focus on a more tractable Hamiltonian

$$H = e^{-\delta t} F\, U(C/F) + \lambda S$$

where we have now just one dimension to the problem.

We now return to the specification of the annuities.

Using the isoelastic utility function

$$U = \frac{A}{\beta}\,(C/F)^{\beta} \tag{6.12}$$

the Hamiltonian becomes

$$H = F^{1-\beta} C^{\beta} e^{-\delta t} + \lambda S \tag{6.13}$$

where time subscripts have been omitted for convenience. The necessary conditions for an interior solution are

$$-F^{1-\beta}\beta C^{\beta-1}e^{-\delta t} + \lambda = 0 \tag{6.14}$$

$$F^{1-\beta}\beta C^{\beta-1}e^{-\delta t}\, r = -\dot{\lambda} \tag{6.15}$$

where r is the return on capital. Differentiating (6.14) with respect to time and substituting into (6.15) for $\dot{\lambda}$, we obtain a differential equation defining an optimal path for consumption.

$$\dot{C}/C = (r-\delta)/(1-\beta) + \dot{F}/F \tag{6.16}$$

Integrating and solving for the constant of integration, we obtain

$$c_t^* = c_o e^{gt} \tag{6.17}$$

where $g = (r-\delta)/(1-\beta)$ and $c = (C/F)$.

A numerical solution for c_o can be obtained by using the fact the lifetime wealth equals lifetime consumption, i.e.

$$W_o = \int_0^T c_t F_t\, e^{-rt}\, dt \tag{6.18}$$

which yields a solution

$$c_o = W_o [\int_0^T F_t\, e^{(g-r)t}\, dt]^{-1} \tag{6.19}$$

To derive a constant consumption value which would yield the same lifetime utility as the optimal consumption path, we must solve

$$\int_0^T \frac{A}{\beta} F_t(\overline{C}/F_t)^{\beta} e^{-\delta t}\, dt = \int_0^T \frac{A}{\beta} F_t\, c_t^{*\beta}\, e^{-\delta t}\, dt \tag{6.20}$$

for $\overline{C}$. This can be done by substituting for c* using (6.17) and in turn for c_o by using (6.19). This yields

$$\int_0^T F_t^{1-\beta}\, \overline{C}^{\beta} e^{-\delta t}\, dt = k_1 \int_0^T F_t e^{(g\beta-\delta)t}\, dt \tag{6.21}$$

where

$$k_1 = W_o^{\beta}\, [\int_0^T F_t e^{(g-r)t}\, dt]^{-\beta} \tag{6.22}$$

The right hand side of (6.21) can be estimated numerically for each family in discrete terms to yield a scaler. The solution for the constant consumption annuity is thus

$$\overline{C} = \frac{W_o}{\int_0^T F_t e^{(g-r)t}\, dt} \left[\frac{\int_0^T F_t e^{(g\beta-\delta)t}\, dt}{\int_0^T F_t^{1-\beta} e^{-\delta t}\, dt}\right]^{1/\beta} \tag{6.23}$$

The annuity in the case where the utility functional is that in (6.11) is given by

$$\overline{C} = W_o [\int_0^T F_t^b\, e^{(g-r)t}\, dt]^{-1} \left[\frac{\int_0^T F_t^b\, e^{(g\beta-\delta)t} dt}{\int_0^T F_t^{-\beta} e^{-\delta t}\, dt}\right]^{1/\beta} \tag{6.24}$$

where $b = -\beta/(1-\beta)$ and $g = (r-\delta)/(1-\beta)$. The derivation of this result is relegated to Appendix H.

We shall estimate annuities based on each form both because the results may be sensitive to the particular function and because there does not exist universal agreement in the literature (Arrow and Kurz (1970) discuss the issue) regarding which form is appropriate.

6.4 Net Worth Estimates

To derive estimates of lifetime purchasing power, we must consider not just lifetime earnings but also the net asset position of the households.

The accurate estimation of net worth is considerably more difficult than the estimation of lifetime earnings due to the availability of data (or lack thereof) on wealth. The source of the data[5] which enabled us to estimate earnings does not contain sufficient information on wealth to be of any value. (The reasons why these data are inadequate are discussed in Appendix I.) It has thus been necessary to obtain data on wealth holdings from an alternative source. Statistics Canada kindly supplied estimates based on the 1970 Survey of Consumer Finances. The estimates were derived in the following way.

All families in the sample were classified into given age-income groups. For each such group, mean wealth and the standard deviation of the distribution within each group were estimated. Thus, we have an estimate of the life-cycle earnings for each family of a given type plus an estimate of the mean net worth of families of a similar type drawn from another data base (the 1970 Survey). The estimate of net worth is then added to the discounted lifetime earnings of the households to yield lifetime purchasing power.[6]

In doing this, we of course encounter the same problem which exists in the estimation of the earnings. That is, if we attribute the same net worth figure to all households of a given type, this will result in an understatement of the degree of inequality. Thus, we follow the same procedure as in Chapters 4 and 5. Random numbers, generated subject to the appropriate degree of variation, are added to the estimated net worth figures in each cell.[7] This is possible since Statistics Canada also supplied the variance associated with the mean for each household type.

There is one remaining point relating to the wealth estimates. This concerns missing data. In a few instances, the samples which Statistics Canada used (drawn from the 1970 Survey) were not sufficiently large to enable a reliable estimate to be made of the mean and variance of net worth for households of a given type. In such cases (and relatively these were very few), the assumption was made that the mean and variance for households of a given type can be approximated by those statistics in an adjoining cell. This assumption will probably bias the results. But we know the direction of the bias because the adjoining cells used for such purposes define the statistics for a lower age or lower income group. Since the mean and variance of net worth increase with both income and age and since the missing data are for older families with high incomes, we know that we are imputing a figure for net worth which makes the estimates of lifetime purchasing power for these families closer to the mean than is actually the case.[8]

We show in the next chapter that this procedure is consistent with decisions of a similar tvpe at earlier stages of the thesis to the extent that where estimation biases exist they work in the same direction.

NOTES ON CHAPTER SIX

1. The reason for taking the expected value is that for finite samples there will exist a distribution of variances for the variable $(\hat{\gamma} + \eta)$. By taking the exprected value, we insure that the mean of this distribution corresponds to the observed variance. Given the large number of observations however, the sampling distribution is likely to be concentrated around the mean.

2. It is worth noting that this formutation of the addition of the random component implies that all residual variation in earnings is purely transitory. If longitudinal data were available for the individuals in the sample it would be possible to test for the existence of a non stochastic unobservable component in the residuals. It is unfortunate that longitudinal data do not exist as Lillard (1977), using U.S. longitudinal data, has found that a systematic (unobservable) component is present.

3. If we examine historical trends in the growth of real GNP and population, an estimate of 2% per capita growth is reasonable.

4. A more complete treatment of this problem is given in Irvine (1978).

5. Statistics Canada: Incomes 1971 (Census families).

6. Since the net worth estimates are based on 1970 data and the incomes on 1971 data, the net worth data have been boosted by 5% in real terms to make the two data sets comparable.

7. The numbers are generated in the same fashion as for the earnings estimates. Initially, the numbers (γ_i) are distributed uniformly on the interval (0, 1). They are then transformed by the factor $(2\gamma-1)$ so that they have zero mean and lie on the interval $(-1,+1)$. Since the variance of this variable is 1/3 the numbers are scaled by $\sqrt{3}$ so that they have unit variance. Thus $\gamma = 3(2-1)$ is the random variable with zero mean and unit variance. In order that the numbers have the same variance as that which exists in the data, they are further scaled by the standard deviation of the sample data. This latter has been supplied by Statistics Canada.

8. An alternative approach to the missing data problem would be to fit a function to the mean and variance of net worth for households by age and income and estimate a mean and variance. However, such a procedure would be very dubious given the likely severe non-linearities involved and the limited degrees of freedom. Hence, we have opted for an approach whose biases are known.

CHAPTER 7

RESULTS

Using the procedures described in the previous chapters, the annuities have been estimated. The resulting dollar amounts define the expenditure per time period which would yield an equivalent level of family lifetime utility to that defined by the optimal consumption paths in (6.23) and (6.24) and which at the same time would exhaust lifetime wealth.

The purpose of this chapter is to examine the behaviour of the annuity distributions and to compare them with the results obtained in other studies using different methodological approaches. We shall consider both positive and (explicitly) normative summary measures of the degree of inequality and also examine the behaviour of components of the distributions. A fairly complete analysis of inequality statistics has been offered in several works[1] and will not be repeated here. But a brief discussion of each of the measures is included with the results.

The major results are contained in Tables 7.1 through 7.7. Tables 7.1 and 7.2 define inequality measures corresponding to the utility functions U = U(C/F) and U = FU(C/F) respectively; i.e. they define the distribution of the variables in equations (6.24) and (6.23). Tables 7.3, 7.4 and 7.5 define the conventional measures - i.e. the distribution of gross incomes and gross earnings for family heads at a point in time. Table 7.6 defines the annuity distribution for lifecycle earnings only, i.e. it does not include the net asset position of the households. Table 7.7 contains results for the distribution of the Weisbrod-Hansen variable. Tables 7.8 and 7.9 contain results for different definitions of family. We will first examine how these distributions differ. We will then compare them with the results obtained by other researchers. Finally, we will show that the conclusions are robust with respect to the assumptions made during estimation.

7.1 The Results

We note at the outset that the inequality statistics obtained using the utility function U=U(C/F) are very similar to those obtained using the function U=FU(C/F). Thus, our conclusions are robust with respect to the choice of welfare function and possible problems foreseen in Chapter 6 fortunately do not arise.[2] For simplicity, we will refer only to the results corresponding to the function U=FU(C/F). (We use this rather than U=U(C/F) as some authors have expressed strong preferences on the matter.[3])

The summary measures of inequality in Tables 7.2 through 7.5 indicate that there is a lower degree of inequality associated with the distribution of annuities than with the distribution of the income of the head at a point in time and this latter in turn exhibits less inequality than the distribution of earnings.

The relative mean deviation (RMD) is defined as the sum of all absolute deviations from the mean divided by total income. It decreases from a value of .5125 for earnings to .4043 for the annuities. As a measure of inequality, it is not very satisfactory though because it does not satisfy the Pigou-Dalton criterion, i.e. a transfer between two individuals will not be reflected in a change in the RMD unless they lie on different sides of the mean.[4]

The variance does satisfy the Pigou-Dalton criterion. Here, the variance for the annuities exceeds the variance for the earnings and for the incomes. This may be misleading however, since the annuity distribution has the largest mean. The coefficient of variation, unlike the variance, is invariant to an equal proportional change in all values and the estimates indicate that the variation is less for the annuities than either the incomes or earnings.[5]

While these measures are useful, one could argue that it is desirable that a measure attribute more weight to a transfer at the lower end of the distribution than at the upper end. That is, if we redistribute a certain amount from a person with an income of $5,000 to a person with $1,000 it might be desirable that the inequality measure should show a bigger change in the direction of equality than if the redistribution were from a person with an income of $105,000 to one with $101,000. One way of achieving this is to take the logarithm of the incomes and use the standard deviation of the transformation. We see that this measure indicates a substantial reduction in inequality at the lower tail of the distribution when we consider annuities rather than incomes or earnings. The value of the standard deviation of the logs declines from 2.89 for earnings to .61 for annuities. Here then we have quite a strong conclusion: If we are averse to inequality at the lower tail then observed incomes or earnings may be misleading indicators of social welfare. We shall see that this result is corroborated when we examine decile shares.

Theil's entropy measure also indicates that there is a general reduction in inequality when we consider annuities rather than incomes or earnings.[6]

The above statistics are all based on the difference between individual values and the mean. The Gini index and the decile shares on the other hand focus on the relationship between every pair of values and thus on the whole distribution. The Gini index for the earnings distribution is .375, for the income distributions it is .357 and .328 and for the annuity distribution, it is .283. Corresponding to each of these, the cumulative distributions indicate that their Lorenz curves do not intersect.[7] We note also that the biggest difference between the distributions exists in the bottom quintile and the top decile. We shall examine the reason for this presently.

Finally, Tables 7.2 through 7.5 contain results for explicitly normative measures of inequality. At one time, the measures examined above would have been referred to as 'positive' measures of income inequality as it was thought that they did not imply any form of social welfare function. Recent work however has shown that the use of these measures does imply some notion of social welfare. For example, Sheshinski (1974) has shown that a social welfare function of the type

$$S = H\ \{\Sigma_i \Sigma_j\ U^{-1}[\mathrm{Min}\ (U(I_i),\ U(I_j))]\}$$

for any function U and any arbitrary increasing function H can be expressed as a function of average income and the Gini coefficient alone. (I_j is the income of person j). In a similar vein, Atkinson (1970) showed that if social welfare is the sum of individual U functions (which are strictly concave in incomes) and if the Lorenz curve for one distribution (A) lies completely inside that of another (B) then, regardless of the form of U, we can say that the welfare associated with (A) exceeds that associated with (B). Sen (1973) has since shown that his results can be generalized. Likewise, if individual welfare functions are proportional to $x_i \log(1/x_i)$ then Theil's measure could fit the utilitarian framework.

If it is the case then that the 'positive' measures are in fact normative, there is a strong case for specifying a priori the assumptions or properties which we would like the inequality measure to have.

Dalton (1920) proposed that we define social welfare as the sum of actual utilities divided by the total utility which would accrue if everyone had the same income. Atkinson (1970) showed that this is not invariant with respect to positive linear transformations of the utility functions and proposed instead the following.

Define Y_e as the equally distributed equivalent income i.e.

$$Y_e = Y \mid (nU(Y) = \Sigma_i U(Y_i)) \tag{7.1}$$

where n is the number of people in the population. His inequality measure is

$$A = 1 - Y_e/\mu \tag{7.2}$$

where μ is the mean. If the equally distributed measure A is to be invariant with respect to proportional shifts in the income distribution then U must exhibit constant relative risk aversion i.e. U equals $C_1 + C_2 Y^{1-\varepsilon}/1-\varepsilon$. Equation (7.2) then becomes

$$A = 1 - [\Sigma_i \left(\frac{Y_i}{\mu}\right)^{1-\varepsilon} \frac{1}{n}]^{1/(1-\varepsilon)} \tag{7.3}$$

TABLE 7.1

INEQUALITY STATISTICS FOR ANNUITIES USING U=U(C/F)

No. of Observations					659					
No. of Weighted Observations					19455					
Mean					12184.23					
Relative Mean Deviation					.4124					
Variance					39420327.4					
Coefficient of Variation					.5153					
Standard Deviation of Logs					.6174					
Theil's Entropy Measure					.1333					
Gini Index					.2878					
Decile Shares	.0265	.0453	.0616	.0752	.0877	.0986	.1150	.1320	.1562	.2020
Cumulative	.0265	.0718	.1334	.2085	.2963	.3949	.5099	.6418	.7980	1.0000
(First 9 deciles each form 9.997 percent of total)										
Atkinson's Measure (Epsilon = .49)					.0676					
Atkinson's Measure (Epsilon = .99)					.1415					
Atkinson's Measure (Epsilon =1.49)					.2203					
Atkinson's Measure (Epsilon =1.99)					.3024					

* These results are based on a random sample from the Statistics Canada Microdata file 'Incomes 1971 (census families)'.

TABLE 7.2

INEQUALITY STATISTICS FOR ANNUITIES USING U=FU(C/F)

No. of Observations					659					
No. of Weighted Observations					19455					
Mean					12571.05					
Relative Mean Deviation					.4043					
Variance					40521890.3					
Coefficient of Variation					.5064					
Standard Deviation of Logs					.6065					
Theil's Entropy Measure					.1290					
Gini Index					.2832					
Decile Shares	.0268	.0463	.0626	.0758	.0881	.0994	.1143	.1306	.1566	.1996
Cumulative	.0268	.0731	.1357	.2115	.2996	.3990	.5132	.6438	.8004	1.0000
(First 9 deciles each form 9.997 percent of total)										
Atkinson's Measure (Epsilon = .49)					.0655					
Atkinson's Measure (Epsilon = .99)					.1372					
Atkinson's Measure (Epsilon =1.49)					.2139					
Atkinson's Measure (Epsilon =1.99)					.2945					

* These results are based on a random sample from the Statistics Canada Microdata file 'Incomes 1971 (census families)'.

TABLE 7.3

INEQUALITY STATISTICS FOR GROSS INCOME OF HEAD WHERE INCOME INCLUDES GOVERNMENT TRANSFERS

No. of Observations	659
No. of Weighted Observations	19455
Mean	8032.07
Relative Mean Deviation	.4486
Variance	27892030.0
Coefficient of Variation	.6575
Standard Deviation of Logs	1.2376
Theil's Entropy Measure	.1956
Gini Index	.3275

Decile Shares	.0135	.0378	.0609	.0774	.0881	.0992	.1137	.1276	.1476	.2342
Cumulative	.0135	.0513	.1122	.1896	.2777	.3768	.4905	.6181	.7658	1.0000

(First 9 deciles each form 9.997 percent of total)

Atkinson's Measure (Epsilon = .49)	.1027
Atkinson's Measure (Epsilon = .99)	.2641
Atkinson's Measure (Epsilon =1.49)	.7803
Atkinson's Measure (Epsilon =1.99)	.9884

* These results are based on a random sample from the Statistics Canada Microdata file 'Incomes 1971 (census families)'.

TABLE 7.4

INEQUALITY STATISTICS FOR GROSS INCOME OF HEAD, WHERE INCOME DOES NOT INCLUDE GOVERNMENT TRANSFERS

No. of Observations	659									
No. of Weighted Observations	19455									
Mean	7689.11									
Relative Mean Deviation	.4872									
Variance	29513872.8									
Coefficient of Variation	.7065									
Standard Deviation of Logs	2.2772									
Theil's Entropy Measure	.2441									
Gini Index	.3566									
Decile Shares	.0029	.0300	.0585	.0772	.0890	.1012	.1168	.1308	.1515	.2422
Cumulative	.0029	.0328	.0913	.1685	.2575	.3587	.4755	.6063	.7578	1.0000
(First 9 deciles each form 9.997 percent of total)										
Atkinson's Measure (Epsilon = .49)	.1442									
Atkinson's Measure (Epsilon = .99)	.4829									
Atkinson's Measure (Epsilon =1.49)	.9701									
Atkinson's Measure (Epsilon =1.99)	.9976									

* These results are based on a random sample from the Statistics Canada Microdata file 'Incomes 1971 (census families)'.

TABLE 7.5

INEQUALITY STATISTICS FOR EARNINGS OF HEAD

No. of Observations	659									
No. of Weighted Observations	19455									
Mean	7364.91									
Relative Mean Deviation	.5125									
Variance	29864898.8									
Coefficient of Variation	.7420									
Standard Deviation of Logs	2.8913									
Theil's Entropy Measure	.2775									
Gini Index	.3750									
Decile Shares	.0000	.0210	.0559	.0771	.0902	.1029	.1179	.1336	.1548	.2466
Cumulative	.0000	.0210	.0770	.1541	.2443	.3472	.4650	.5986	.7534	1.0000
(First 9 deciles each form 9.997 percent of total)										
Atkinson's Measure (Epsilon = .49)	.1731									
Atkinson's Measure (Epsilon = .99)	.6121									
Atkinson's Measure (Epsilon =1.49)	.9880									
Atkinson's Measure (Epsilon =1.99)	.9986									

* These results are based on a random sample from the Statistics Canada Microdata file 'Incomes 1971 (census families)'.

Small values of ε indicate indifference to inequality ($\varepsilon = 0$ yields $A = 0$) and high values of ε mean that society is averse to inequality.[8] The maximum value for A is unity. Thus, in Table 7.2 $A = .1373$ for $\varepsilon = .99$. This means that if our values are such that we believe ε should equal .99 then society would need only 87.27% of its present total 'income' to attain the same degree of social welfare. It is uniformly the case that the statistics follow the same pattern as the previous 'positive' measures. The choice of the appropriate yardstick depends on the value judgements made.

To conclude this section, we could summarize by saying that the results indicate that the 'true' (i.e. annuity) distribution defines a lower level of inequality than the income or earnings distributions would indicate by almost every criterion. We now analyze <u>why</u> we have obtained such results and we shall show that the above statistics form a lower bound to the real degree of inequality.

7.2 Analysis of Results

The focus of our attention in this section is on an examination of why the annuity distribution on the one hand differs from the earnings and income distributions on the other. The reason for the difference between the earnings and income distributions is straightforward. The income variable includes transfer payments and retirement income and thus many households who have zero earnings do not have zero income. (We note that the lowest decile in the earnings distribution has zero percent of total earnings while the lowest income decile has 1.35% of total income - Table 7.3.)

In analysing why we obtain an apparent reduction in inequality when we estimate the annuity distribution, an obvious approach is to decompose the variation in lifetime purchasing power; i.e. is it greater equality in life-cycle earnings or greater equality in wealth holdings or just greater equality in the <u>joint</u> distribution of lifetime earnings and wealth which accounts for the apparent equality?

Accordingly, the inequality measures have been reestimated using only life-cycle earnings. The estimates are presented in Table 7.6. As the results are invariant to the form of welfare function used, we present only those based on the function U = FU(C/F).

The annuity distribution based on life-cycle earnings alone exhibits greater inequality than that which includes net worth in addition. While this distribution is less unequal than the distribution of observed earnings and the distribution of incomes excluding transfers, it exhibits a degree of inequality very close to that displayed by the distribution of incomes including transfers. This is

TABLE 7.6

INEQUALITY STATISTICS FOR ANNUITIES USING U=FU(C/F) WHERE ANNUITIES ARE ESTIMATED ON THE BASIS OF LIFE-CYCLE EARNINGS ALONE

No. of Observations	659									
No. of Weighted Observations	19455									
Mean	11036.69									
Relative Mean Deviation	.4775									
Variance	42857070.2									
Coefficient of Variation	.5932									
Standard Deviation of Logs	.9056									
Theil's Entropy Measure	.1825									
Gini Index	.3326									
Decile Shares	.0153	.0357	.0557	.0700	.0861	.1006	.1175	.1367	.1662	.2162
Cumulative	.0153	.0510	.1067	.1768	.2629	.3636	.4810	.6177	.7838	1.0000
(First 9 deciles each form 9.997 percent of total)										
Atkinson's Measure (Epsilon = .49)	.0962									
Atkinson's Measure (Epsilon = .99)	.2174									
Atkinson's Measure (Epsilon =1.49)	.4883									
Atkinson's Measure (Epsilon =1.99)	.9503									

* These results are based on a random sample from the Statistics Canada Microdata file 'Incomes 1971 (census families)'.

interesting because it indicates that life-cycle factors do indeed even out some of the inequality in the distribution of earnings.

In comparing the annuitised earnings distribution with the distribution of incomes net of transfers, we see that by every criterion the former displays more equality than the latter. Their Lorenz curves are drawn in Figure 7.1. We can therefore conclude that if the distribution of net incomes at a point in time is used by policy makers as a measure of inequality, it yields an unduly pessimistic view. This is particularly true if we are concerned with the lower tail.

If we now compare the distribution of annuitised earnings with the distribution of incomes including transfers, we find it more difficult to make statements regarding relative inequality. If we were simply to examine the Gini coefficients we would conclude that the income distribution displays less inequality than the annuitised earnings distribution. However, if we examine the cumulative distributions, we see that the Lorenz curves intersect <u>twice</u> (see Figure 7.2). Thus, the effect of considering life-cycle earnings is to even out the distribution at the tails and have a 'middle heavy' distribution.

This phenomenon is attributable to the behaviour of the

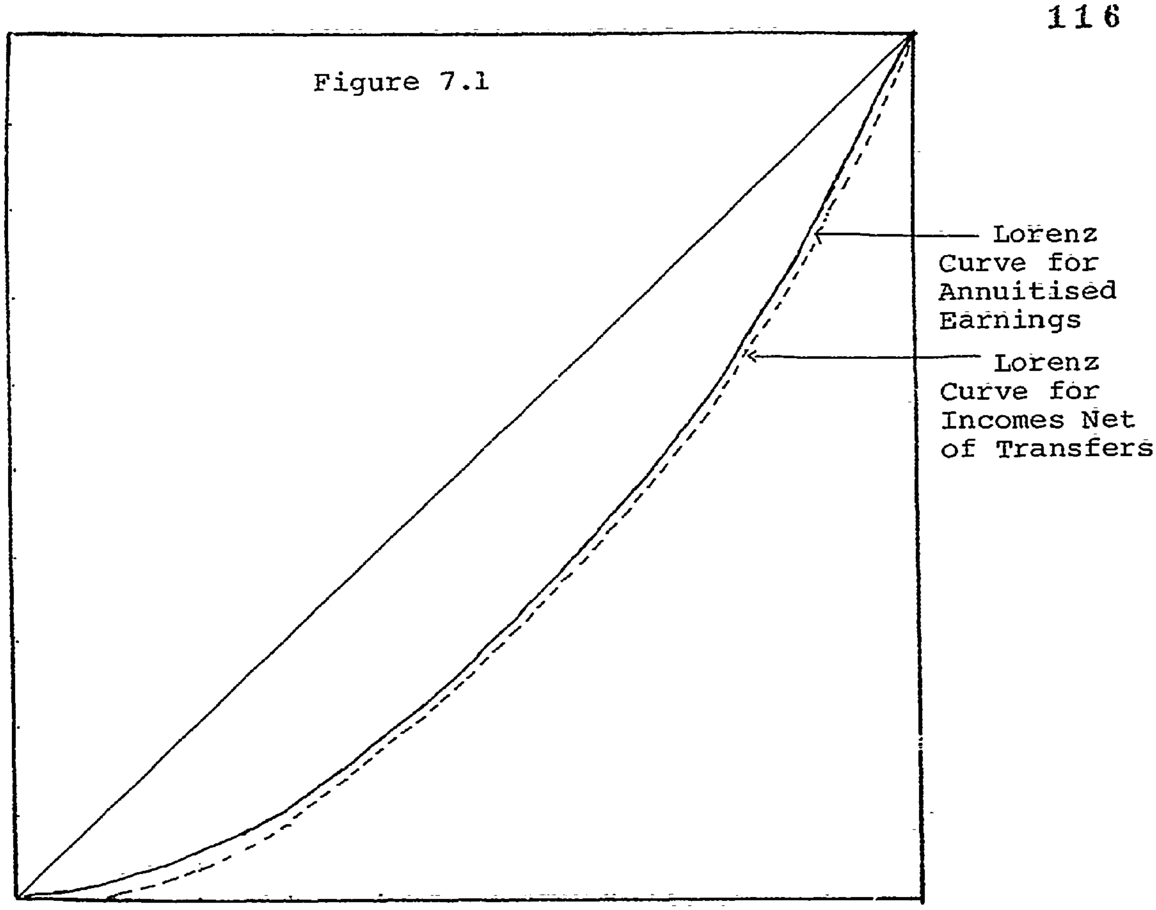


Figure 7.1

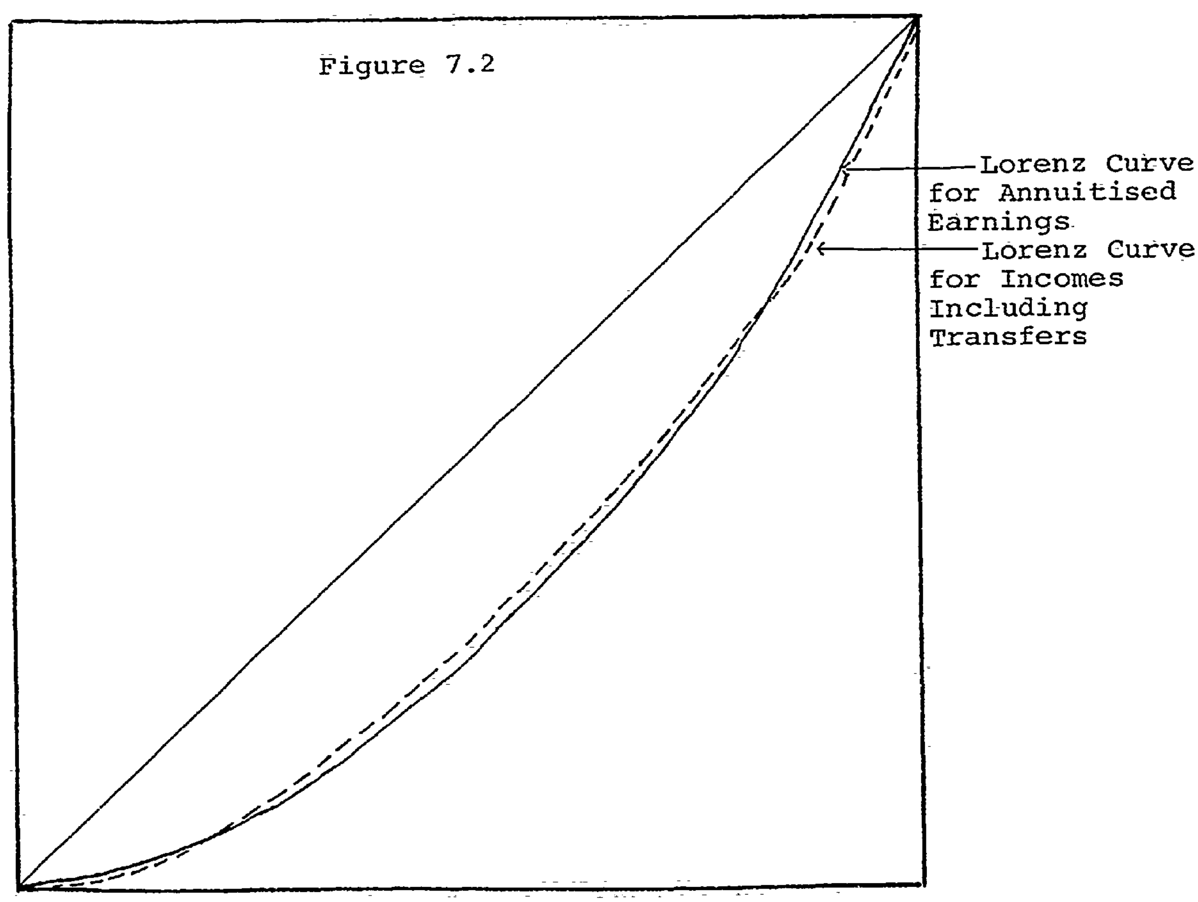


Figure 7.2

transitory component of income. Many of the individuals in the bottom quintile did not work in the survey year (for at least part of the year). As a consequence, their earnings and income in that year are low, and government transfers form a large part of their income. However, their unemployment is not permanent and over the course of their lifetime, their average earnings exceed the observed income in that particular year. The reason for the evening out in the upper tail is that some individuals with high incomes have large positive transitory components. While the probability of unemployment for this group is very low, the fact that positive transitory components exist means that average earnings over the course of the life-cycle will be less.

An analysis of the employment behaviour of household heads in the base year yields the following information. Of the top quintile (of incomes) 93.1% worked full time (in terms of weeks per year) while only 35.1% of the bottom quintile worked full time.[9] The average for the population was 73.6%. A similar picture emerges if we consider full time workers in terms of hours per week rather than weeks per year. The figures are 98.5%, 51.1% and 86.5% respectively.

The fact that the Lorenz curves intersect twice is important for several reasons. First, it shows the inadequacy of most summary measures of inequality because they conceal what may be taking place throughout the range of the distribution. Second, even though the Gini index for the observed income distribution is less than that for the earnings-annuity distribution, an explicitly normative evaluation may indicate that there is a greater degree of equality associated with the earnings-annuity distribution due to the evening out at the tails. We see from Tables 7.3 and 7.6 that Atkinson's measures indicate this. Finally, it is important because it illustrates the importance of dynamic rather than purely static measures of inequality. Since the greatest concern of policy makers is with the lower tail of the distribution, policy makers can take some satisfaction from the fact that the plight of all members of this group is not a permanent one.

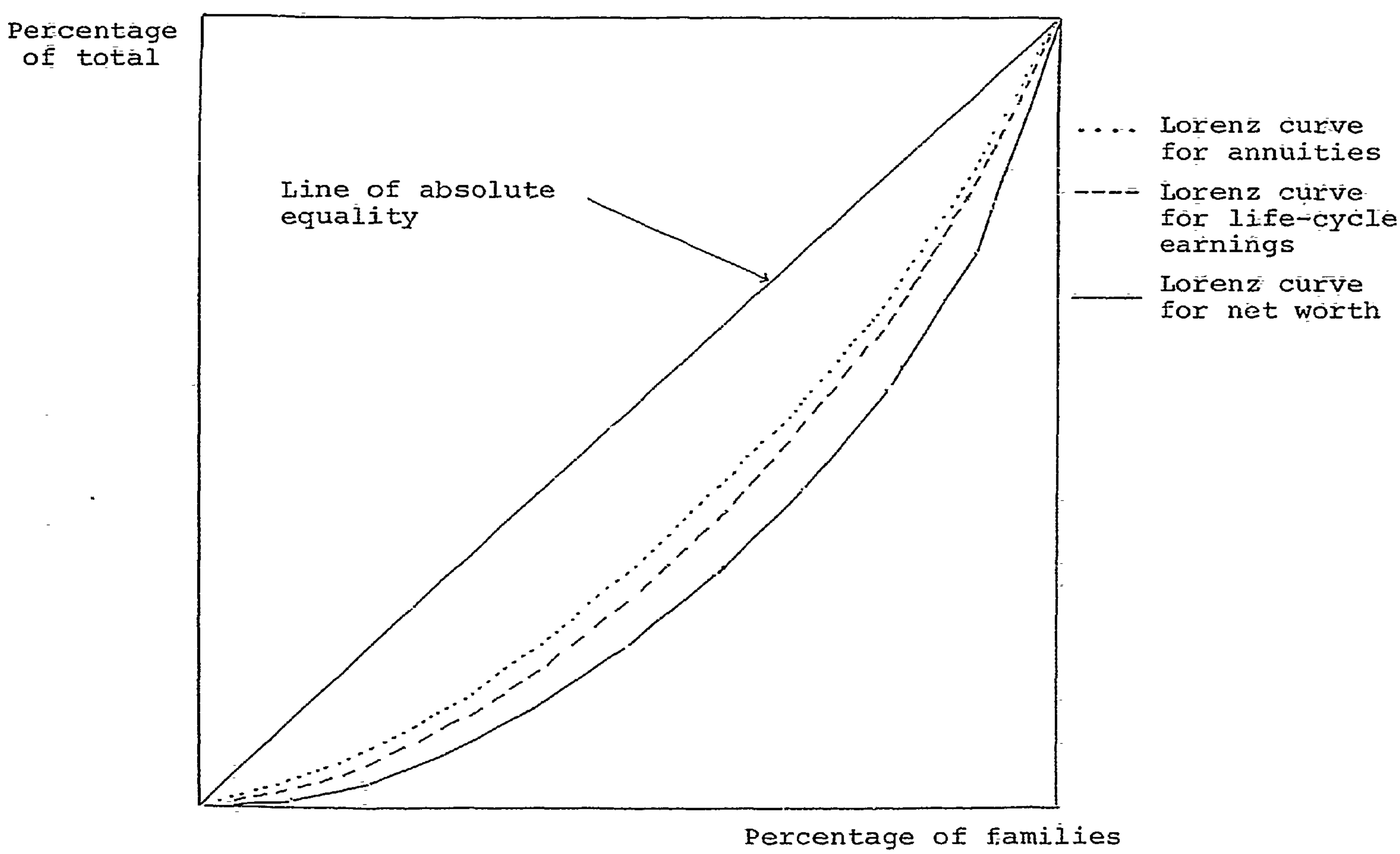
In light of the above, it seems that the difference between the income distribution and the annuity distribution is primarily attributable to the manner in which wealth is distributed. It is important to point out though that this does not imply that net worth is more evenly distributed than life-cycle earnings. It is possible that the Lorenz curve for net worth could lie completely outside the Lorenz curve for lifetime earnings while the Lorenz curve for the annuities could lie inside both.[10] The reason is to be found in an examination of the <u>joint</u> distribution of lifetime earnings and net worth. If low income families have large stocks of assets, the annuity distribution will exhibit more equality than either of its components. These data indicate that there is a certain amount of 'evening-out' taking place (see Figure 7.3).

As this is a very important issue - indeed it formed the major hypothesis proposed in Chapter 1 - it is necessary (a) to examine if the above corresponds to the results obtained by other researchers and (b) to examine the effects of known biases which were introduced, of necessity, during the estimation.

The methodological approach used in this thesis has not been utilized by other researchers using Canadian data. However, several researchers have examined the distribution of income in conjunction with the distribution of wealth. Podoluk (1974), Wolfson (1977) and Beach and Flatters (1978) have all used micro data.

Podoluk concluded that wealth is less evenly distributed than income, but is more evenly distributed between income groups than is income (Figure 2.1). Both Wolfson and Beach and Flatters examined the joint distribution of income and wealth by utilizing an approach developed by Weisbrod and Hansen (1968). These latter estimated an annuity based on the net worth of families and analysed the distribution of a variable which was the sum of observed income at a point

FIGURE 7.3


Percentage of total
Line of absolute equality
Lorenz curve for annuities
Lorenz curve for life-cycle earnings
Lorenz curve for net worth
Percentage of families

in time plus this annuity.

Wolfson found that the distribution of the Hansen-Weisbrod variable indicated approximately the same degree of inequality as the distribution of income alone. With an interest rate of 4%, the Gini index decreased by 3% and with a 10% rate, the index increased by 2%. (In contrast, Weisbrod and Hansen found that inequality increased markedly.)

Beach and Flatters experienced data problems similar to those encountered in this study.[11] That is, they had to estimate net worth for the households in their sample. Using discount rates of 8% and 12%, they found that inequality increased with the inclusion of the annuity, and by a greater amount for the 12% rate.

In the light of these results, it may seem a little surprising that this study indicates that inequality decreases somewhat when net worth is accounted for. But this is not the case. We stated in the last section that the results form a lower bound to the actual degree of inequality. This is because we did not have estimates of net worth for some groups who were both old and had a high income. We assumed, for lack of a better alternative, that the wealth holdings would be the same for these families as a group with a lower income. This, of course, affects the results, as the distribution of wealth is known to be skewed in the upper tail.

Effectively, our estimation procedure has cut off the tail and it is thus not surprising that our results are biased in a downward direction.

As a means of checking this postulate, we estimated the Weisbrod-Hansen annuities (using a discount rate of 5%). The results are given in Table 7.7. We see that this distribution indicates a slightly lower degree of inequality than the distribution of incomes. This is consistent with the proposition that our estimates in Table 7.2 form a lower bound to the true degree of inequality.[12]

There are two further reasons for believing that we have been conservative in defining the degree of inequality.

First, it is possible that the estimates of net worth from the Survey of Consumer Finances are understated.[13] Second, the earnings equations in Chapters 4 and 5 may not be fully specified. Earnings were estimated as a function of a set of observable variables plus a random term. The random term could be interpreted as transitory income. The approach, however, assumes that for two like individuals, each has the same probability of having a positive or negative transitory income. It may be the case though that we could obtain a better estimate of their future earnings if we had some sort of an achievement index for the individuals. The fact that certain unobservable variables may be relevant

TABLE 7.7

INEQUALITY STATISTICS FOR WEISBROD-HANSEN ANNUITIES

No. of Observations	659									
No. of Weighted Observations	19455									
Mean	9478.6									
Relative Mean Deviation	.4442									
Variance	37789171.8									
Coefficient of Variation	.6485									
Standard Deviation of Logs	.9301									
Theil's Entropy Measure	.1858									
Gini Index	.3229									
Decile Shares	.0182	.0429	.0606	.0746	.0851	.0964	.1099	.1259	.1470	.2393
Cumulative	.0182	.0612	.1218	.1963	.2815	.3779	.4878	.6137	.7607	1.000
(First 9 deciles each form 9.997% of total)										
Atkinson's Measure (Epsilon = .49)	.0942									
Atkinson's Measure (Epsilon = .99)	.2153									
Atkinson's Measure (Epsilon =1.49)	.4991									
Atkinson's Measure (Epsilon =1.99)	.9408									

* These results are based on a random sample from the Statistics Canada Microdata file 'Incomes 1971 (census families)'.

means that we may again be cutting down on the degree of variation in earnings. If there existed longitudinal data, it would be possible to estimate, for each individual in the sample, whether the random component of earnings should be truly random (in the sense of having an equal probability of being positive or negative) or whether the probability of it being positive should be different from the probability of it being negative. It is thus unfortunate that no such data exist for Canada.

At this stage it would be useful to summarise the thrust of the analysis thus far.

We started out by postulating that the distributional measure implicit in the model of lifecycle utility maximization forms a more meaningful basis for distributional analysis than the conventional measures. Accordingly, we have compared the distribution of annuities with several distributions of the conventional type. The annuity distribution, however, is based on a model of <u>family</u> optimization while the other distributions we have examined are based on the income or earnings of family heads alone. This procedure has been deliberate. For the analysis has attempted to focus upon how inappropriate distributional measures could lead to the wrong type of conclusion when compared with the 'correct' measure (i.e. the annuities). The fact that the conventional measures use individual earnings or incomes while the model developed in this thesis uses family earnings is a byproduct of the focus of the analysis. Nonetheless, it is valid and interesting to examine how a distribution of annuities based on net worth plus the lifecycle earnings of the family head alone would compare with the distributions which have been examined to date. This procedure would complete the analysis. For in this chapter we have been able to identify several reasons for the differences which exist between the distributions. The only issue we have not examined in detail is that of family size. We now do this.

Table 7.8 contains inequality statistics for the distribution of annuities based on net worth plus lifecycle earnings of male household heads. The annuity formulae (equations (6.23) and (6.24)) in this case are simply defined by equation (3.17) since $F_t = 1$ in every time period, and the question of different utility functions does not arise.

We see from table 7.8 that the distribution of these annuities is slightly less equitable than the distribution of the annuities based on family earnings. The coefficient of variation is .57 instead of .51, the standard deviation of the logs is .75 instead of .61 and the Gini index is .319 instead of .283. The decile shares indicate that the Lorenz curves do not intersect and Atkinson's measures likewise indicate that no reversals take place for different values of the inequality aversion parameter.

TABLE 7.8

INEQUALITY STATISTICS FOR ANNUITISED VALUE OF LIFECYCLE EARNINGS AND NET WORTH OF INDIVIDUAL MALE HOUSEHOLD HEADS

No. of Observations	659									
No. of Weighted Observations	19455									
Mean	10839.1									
Relative Mean Deviation	.4591									
Variance	38191337.1									
Coefficient of Variation	.5702									
Standard Deviation of Logs	.7480									
Theil's Entropy Measure	.1664									
Gini Index	.3189									
Decile Shares	.0183	.0385	.0581	.0725	.0854	.0992	.1178	.1372	.1610	.2119
Cumulative	.0183	.0569	.1149	.1874	.2728	.3720	.4898	.6271	.7881	1.000
(First 9 deciles for 9.997% of total)										
Atkinson's Measure (Epsilon = .49)	.0862									
Atkinson's Measure (Epsilon = .99)	.1851									
Atkinson's Measure (Epsilon =1.49)	.2954									
Atkinson's Measure (Epsilon =1.99)	.4140									

* These results are based on a random sample from the Statistics Canada Microdata file 'Incomes 1971 (census families)'.

TABLE 7.9

INEQUALITY STATISTICS FOR ANNUITISED VALUE OF LIFECYCLE EARNINGS ALONE OF INDIVIDUAL MALE HOUSEHOLD HEADS

No. of Observations	659
No. of Weighted Observations	19455
Mean	9572.3

Relative Mean Deviation	.5386
Variance	40414667.3
Coefficient if Variation	.6641
Standard Deviation of Logs	1.1667
Theil's Entropy Measure	.2310
Gini Index	.3711

Decile Shares	.0088	.0278	.0493	.0666	.0810	.0991	.1206	.1447	.1709	.2313
Cumulative	.0088	.0366	.0859	.1525	.2335	.3326	.4532	.5979	.7687	1.000

(First 9 decile form 9.997% of total)

Atkinson's Measure	(Epsilon = .49)	.1248
Atkinson's Measure	(Epsilon = .99)	.2937
Atkinson's Measure	(Epsilon =1.49)	.6628
Atkinson's Measure	(Epsilon =1.99)	.9753

* These results are based on a random sample from the Statistics Canada Microdata file 'Incomes 1971 (census families)'.

The important conclusion from this is that the distribution of individual annuities is considerably more equitable than the distribution of individual earnings at a point in time. A comparison of Tables 7.8 and 7.5 shows that this is true by every criterion.

As in the earlier part of this chapter we can again examine why it is that the distribution of individual annuities is more equitable than the distribution of observed earnings. In table 7.9 we present the inequality statistics which define the distribution of individual annuities based on lifecycle earnings alone. We again find that this distribution exhibits greater equality than the distribution of observed earnings - particularly at the tails.

We thus come to the same conclusion in our analysis of individual behaviour as we came to in the analysis of family behaviour. That is: the distribution of annuities based on lifecycle earnings plus net worth is more equitable than the distribution of annuitised lifecycle earnings alone, and this latter is more equitable than the distribution of observed earnings at a point in time.

While Theil's measure is interesting, Sen (1973, p. 36) points out that

> "...the fact remains that it is an arbitrary formula, and the average of the logarithms of the reciprocals of income shares weighted by income shares is not a measure that is exactly overflowing with intuitive sense."

7. If the distributions had the same mean, we could use a theorem proved by Atkinson (1970) to claim that the annuity distribution is preferable to the income distribution which is in turn preferable to the earnings distribution. As they do not have the same mean, we cannot make this claim.

8. If $\varepsilon = 1$ it means that society would consider it fair to take $10 from a rich person and give $5 to a person with half that income. If $\varepsilon = 2$, just $2.5 would have to be given to the person lower in the distribution to leave society at a level of welfare equivalent to that which existed before the redistribution.

9. A full time worker is one who works 50 weeks or more, mostly on a full time basis (in terms of hours per week). A part time worker is one who works less than 50 weeks or works mostly on a part time basis (in terms of hours per week).

10. The data supplied by Statistics Canada yield a Gini index of .4303 (subject to the estimation conditions laid out in Chapter 6, section 6.4). The lowest decile has less than 1% of wealth and the top quintile has 46% of wealth. As we indicated in Chapter 6, these are likely very conservative estimates of the degree of inequality associated with wealth holdings. For example, Podoluk (1974), using the same data base, came up with a substantially larger Gini index. Nonetheless, her results are consistent with those presented in the text.

11. Their approach is outlined in Chapter 9 of their monograph. It bears a likeness to the methodology adopted in this study as allowance was made for the variation in wealth holdings among households of a similar type.

12. This is because previous authors generally found a slight increase in inequality. In this study, we have found a slight decrease (see Tables 7.4 and 7.7). This supports the contention that the manner in which wealth holdings have been estimated biases the results downwards.

13. Davies (1978) provides evidence which supports the widely held belief that underreporting is a serious problem with survey data, particularly in the upper tail of the distribution.

NOTES ON CHAPTER SEVEN

1. See, for example, Sen (1973) or Atkinson (1970).

2. See Chapter 6, section 6.3.

3. See Arrow and Kurz (1970).

4. If a measure indicates a lower degree of inequality as a result of a transfer from a rich person to a poor person then it satisfies the Pigou-Dalton criterion. Dalton (1920) discusses this in detail.

5. It can reasonably be argued however, that a measure of inequality should not be purely relative. For example; in a three person economy incomes of $100, $200 and $300 would yield the same coefficient of variation as incomes of $1,000, $2,000 and $3,000. If the inequality measure enters a social welfare function then a tenfold increase in all incomes would not be reflected as an increase in welfare if we were to use the coefficient of variation. Consequently, a choice between the variance and coefficient of variation must be based on whether one views poverty as a relative or absolute concept.

6. Theil's entropy measure - as developed in Theil (1967) - is given by

$$E = \sum_{i=1}^{n} x_i \log (1/x_i) \qquad (1)$$

If x is the probability that an event will occur and log(1/x) the information content of the event then the entropy of the situation is the sum of the information content of each event weighted by the respective probability. In inequality theory if x_i is the share of total income going to person i, then the maximum possible degree of equality exists when each x_i equals 1/n i.e. E = log(n). Theil's measure is defined as the difference between log(n) and equation (1) above. This can be shown to be

$$T = \sum_{i=1}^{n} x_i \log(nx_i) \qquad (2)$$

CHAPTER 8

SUMMARY AND CONCLUSIONS

At the start of this thesis, two postulates were made regarding distributional studies. First, we stated that such studies should incorporate an analysis of both wealth and earnings since both represent purchasing power, even though one is a stock and the other a flow. Second, we stated that since some part of observed inequality at a point in time can be attributed to systematic changes in earnings over the life-cycle, distributional analysis should seek to incorporate this factor also.

In accordance with these axioms, we estimated lifetime earnings for a large number of families in Canada. To the discounted value of such earnings an estimate of net worth was added. This value of lifetime purchasing power was then annuitised over the course of the life of each family in accordance with a model based on optimal consumption behaviour. This yielded a distribution of purchasing power for a segment of the population. It is proposed that this distribution is more meaningful than the traditional types of distribution (e.g. the distribution of observed incomes, earnings or wealth at a point in time), because it is a better representation of purchasing power in a dynamic framework than these latter.

Three major findings have emerged from this study, each of which has a different implication for economic policy making.

(a) The first conclusion we draw from the study relates to the behaviour of earnings over the lifecycle. The distribution of annuities based on lifecycle earnings alone is considerably more equitable than the distribution of earnings at a point in time. It is also more equitable than the distribution of observed incomes (excluding transfers), though not significantly different from the distribution of incomes inclusive of transfers. The inclusion of net worth in the annuities with lifecycle earnings yields a distribution which in turn is more equal than the distribution of lifecycle earnings. This illustrates the importance of considering the joint distribution of earnings and wealth because, as we illustrated in the thesis, both components of the distribution are less equitably distributed than the joint distribution.

Similar conclusions hold when the distribution of annuities for individual male household heads is examined.

(b) As well as being able to make statements about the bahaviour of the whole distribution, it is important also to examine how the components of the annuity distribution differ from the components of the earnings and income distributions. Despite the fact that summary measures of inequality did not indicate a significant difference existed between the annuitised earnings distribution and the distribution of observed incomes (including transfers), a more detailed analysis indicates that there exists a substantial difference between them if we consider different segments of the distribution. In particular, we showed that the bottom quintile is not as badly off when we consider the annuities rather than current incomes or earnings. This is attributable to the (un)employment behaviour of this group over time. Thus, while particular individuals find themselves at the bottom of the earnings or income distribution in one time period due to the fact that they are unemployed for part of the period, they can expect to move up in the distribution when their employment situation improves. Government policies aimed at this group should thus perhaps consider improving capital markets. They should also be concerned with tracking the behaviour of this group over time to find out if there exists substantial mobility in Canada between income groups. For even though our results indicate that the lowest group (quintile) are not as badly off as traditional indices suggest, an increase in their share of the economic pie from 5.13% to 7.31% would not satisfy the wishes of all social reformers.

(c) Our final conclusion pertains to the upper tail of the distribution and the quality of data now available in Canada. Given the inadequacies of the data we are forced to state that our conclusion regarding the reduction in share accruing to the upper tail is a tentative one. Our policy implications here are very clear. We need much better data before we can formulate economic policies aimed at this group. In future surveys, data will be made available for both income and wealth of each household. This will constitute a considerable improvement over the present situation as we would then have actual rather than estimated wealth figures. If, in addition, longitudinal data on earnings were collected in Canada, the quality of policies geared to redistribution would improve immeasurably.

BIBLIOGRAPHY

Adelman, I. and Morris, C.T. "Performance criteria for evaluating economic development potential - an operational approach". Quarterly Journal of Economics 1968.

Aitchison, J. and Brown, J. "On criteria for descriptions of the income distribution". Metroeconomica 1954.

Andrews, F., Morgan, J. et al. "Multiple classification analysis". University of Michigan 1973.

Applebaum, E. and Harris, R. "Imperfect capital markets and life cycle saving". Canadian Journal of Economics 1978.

Atkinson, A.B. "On the measurement of economic inequality". Journal of Economic Theory 1970.

Ammon, O. "Some social applications of the doctrine of philosophy". Journal of Political Economy 1899.

Arrow, K. and Kurz, M. "Public investment, the rate of return, and optimal fiscal policy". Johns Hopkins 1970.

Baumol, J. "Business behaviour, value and growth". Macmillan 1959.

Beach, C. and Flatters, F. "The income and asset distribution picture in Ontario". Ontario Economic Council 1978.

Becker, G. "Human capital: A theoretical and empirical analysis with special reference to education". Columbia Univeristy Press 1964.

Ben-Porath, Y. "The production function of human capital and lifecycle earnings". Journal of Political Economy 1967.

Blinder, A. "Toward an economic theory of income distribution". M.I.T. Press 1974.

Boissevain, C. "Distribution of abilities depending upon two or more independent factors". Metron 1939.

Bowley, A. "The action of economic forces in producing frequency distributions of incomes, prices and other phenomena". Econometrica 1933.

Cain, G. "Married women in the labour force: An economic analysis". Univeristy of Chicago Press 1966.

Champernowne, D. "A model of income distribution". Economic Journal 1953.

Chiswick, B. "Income inequality: Regional analysis within a human capital framework". N.B.E.R. 1974.

Dagum, C. and Theoret, C. "The impact of the composition of income on the size distribution of income". Research paper no. 7607, Department of Economics, University of Ottawa 1977.

Dalton, H. "The measurement of the inequality of incomes". Economic Journal 1920.

Dasgupta, P., Sen, A. and Starrett, D. "Notes on the measurement of inequality". Journal of Economic Theory 1973.

DaVanzo, DeTray and Greenberg. "Estimating labour supply response: a sensitivity analysis". Office for Economic Opportunity 1973.

Davies, J. "The impact of inheritance on lifetime income inequality in the United States". Mimeo, University of Western Ontario, Department of Economics 1978.

Dodge, D. "Impact of tax, transfer payments and expenditure policies of government on the redistribution of personal income in Canada". Review of Income and Wealth 1975.

Due, J. "Net worth taxation". Public Finance 1960.

Eckaus, El Safty and Norman. "An appraisal of the calculations of rates of return to higher education". Gordon, M.S. ed. "Higher education and the labour market". McGraw-Hill, 1974.

Friedman, M. "Choice, chance and the personal distribution of income". Journal of Political Economy 1953.

Friedman, M. "A theory of the consumption function". Princeton University Press 1953.

Friedman, M. and Kuznets, S. "Income from independent professional practice". N.B.E.R. 1954.

Gibrat, R. "On economic inequalities". International Economic Papers 1957.

Gillespie, I. "On the redistribution of income in Canada". Canadian Tax Journal 1976.

Goldberger, A. "Economic theory". Wiley 1964.

Goldfeld, S. and Quandt, R. "Some tests for homoskedasticity". Journal of the American Statistical Association 1965.

Haldane, J. "Moments of the distribution of powers and products of normal variates". Biometrica 1941/42.

Hanoch, G. "An economic analysis of schooling". Journal of Human Resources 1967.

Henderson, D.W. and Rowely, J. "The distribution and evolution of Canadian family incomes 1965-1973". Economic Council of Canada 1978.

Henle, P. "Exploring the distribution of earned income". Monthly Labour Review, December 1972.

Irvine, I. "Pitfalls in the specification of optimal lifetime consumption patterns". Oxford Economic Papers 1978.

Johnston, J. "Econometric methods". McGraw-Hill 1972.

Kalecki, M. "On the Gibrat distribution". Econometrica 1945.

Kmenta, J. "Elements of econometrics". Macmillan 1971.

Kapteyn. "Skew frequency curves in biology and statistics". 1903.

Kendall. "The advanced theory of statistics". 1948.

Lillard, L. "Inequality: Earnings v's human wealth". American Economic Review 1977.

Love, R. and Wolfson, M. "Income inequality: Statistical methodology and Canadian illustrations". Statistics Canada, Occasional Paper no. 13-559 1976.

Lydall, H. "The distribution of employment incomes". Econometrica 1959.

Lydall, H. "The structure of earnings". Oxford 1968.

Mandelbrot, B. "Stable Paretian random functions and the multiplicative variation in income". Econometrica 1961.

Mayer, T. "The distribution of ability and earnings". Review of Economics and Statistics 1960.

Michael, R. "Education and the derived dema d for children". Journal of Political Economy 1972.

Mincer, J. "Investment in human capital and personal income distribution". Journal of Political Economy 1958.

Mincer, J. "Schooling, experience and earnings". N.B.E.R. 1974.

Mincer, J. and Polachek, S. "Family investments in human capital: earnings of women". Journal of Political Economy 1972.

Naylor, T.H. et al. "Computer simulation techniques". Wiley 1966.

Nordhaus, W. "The effects of inflation on the distribution of economic welfare". Journal of Money, Credit and Banking 1973.

Paglin, M. "The measurement of inequality: a basic revision". American Economic Review 1975.

Pareto, V. "Course d'économie politique". Lausanne 1897.

Pigou, A. "Welfare economics". 1932.

Pigou, A. "Wealth and welfare". Macmillan 1912.

Podoluk, J. "Some comparisons of the Canadian - U.S. income distributions". Review of Income and Wealth 1970.

Podoluk, J. "Measurement of the distribution of wealth in Canada". Review of Income and Wealth 1974.

Projector, D. and Weiss, G. "Survey of financial characteristics of consumers". Washington: Board of Governors of the Federal Reserve System 1966.

Rivlin, A. "Income distribution - can economists help?". American Economic Review, Papers and Proceedings 1975.

Roy, A.D. "The distribution of earnings and of individual output". Economic Journal 1950.

Roy, A.D. "A further statistical note on the distribution of individual output". Economic Journal 1950.

Roy, A.D. "Some thoughts on the distribution of earnings". Oxford Economic Papers 1951.

Ruggles, R. and Ruggles, N. "The anatomy of earnings behaviour". Paper presented at 'Conference on Research in Income and Wealth'. N.B.E.R. 1974.

Rutherford, R. "Income distributions - a new model". Econometrica 1955.

Sheshinski, E. "Relation between a social welfare function and the Gini Index". Journal of Economic Theory 1974.

Simon, H. "The compensation of executives". Sociometry 1957.

Simons, H.C. "Personal income taxation". University of Chicago Press 1938.

Solow, R. "Some long run aspects of the distribution of wage incomes". Econometrica 1951.

Staehle, H. "Ability, wages and income". Review of Economics and Statistics 1943.

Strotz, R. "Myopia and inconsistency in dynamic utility maximization". Review of Economic Studies 1955/6.

Taubman, P. and Wales, T. "Higher education and earnings". McGraw-Hill 1974.

Theil, H. "Principles of Econometrics". Wiley 1971.

Theil, H. "Economics and information theory". North Holland 1967.

Tinbergen, J. "Some remarks on the distribution of labour incomes". International Economic Papers 1951.

Tinbergen, J. "On the theory of income distribution". Weltwirtschaftliches Archiv 1956.

Tinbergen, J. "Welfare economics and income distribution". American Economic Review, Papers and Proceedings 1957.

Tinbergen, J. "A positive and a normative theory of income distribution". Review of Income and Wealth 1970.

Tuckman, H.P. and Brosch, G. "Changes in personal income and their effect on income shares". Southern Economic Journal 1974.

Vickrey, W. "Averaging income for income tax purposes". Journal of Political Economy 1939.

Weisbrod, B. and Hansen, W. "An income net-worth approach to measuring economic welfare". American Economic Review 1968.

Wold, H. and Whittle, P. "A model explaining the Pareto distribution of wealth". Econometrica 1957.

Wolfson, M. "Wealth and the distribution of income, Canada 1969-1970". Paper presented at the 1977 Meetings of the Canadian Economics Association.

Yaari, M. "On the consumer's lifetime allocation process". International Economic Review 1964.

Yaari, M. "Uncertain lifetime, life insurance and the theory of the consumer". Review of Economic Studies 1965.

Statistics Canada. Microdata file: 'Incomes 1971 (census families)', Consumer Income and Expenditure Division, Survey of Consumer Finances 1972.

Statistics Canada. Survey of consumer finances 1970.

APPENDIX A

DATA

The data used in this study came from two sources. The source which serves for the estimation of the earnings profiles is the 1972 Survey of Consumer Finances which used the full April 1972 Labour Force Survey Sample. This is described below under the headings (1) Sample, (2) Weighting and (3) Data Collection. The second data source is the May 1970 Survey of Consumer Finances. This contains information on wealth holdings (which the 1972 survey does not). The process whereby the sources are matched is described in Chapter 6, section 6.4.

A.1 The Sample

The 1972 Survey of Consumer Finances used the full April 1972 Labour Force Survey Sample.

The sample consisted of 35,386 households. Of these, 3,966 were vacant and another 2,978 could not be contacted or could not be interviewed for other reasons including complete refusal of all information. From the occupied households, satisfactory income data were collected for 26,206 census families and 43,039 individuals aged 14 and over who were in receipt of cash income in 1971. The micro-data file contains

25,927 census families containing 78,389 individuals of whom 40,858 are income recipients.

The sample coverage for the Consumer Finance Survey includes almost all individuals residing in private households. The only exclusions are persons who were resident in the Yukon and Northwest Territories, as well as all individuals who resided in institutions, on Indian reservations, or in military camps.

A.2 Weighting

The respondent sampled families from the survey were assigned weights to blow the sample up to national estimates and to compensate for non-response in any particular category of families. The following paragraphs explain how this weighting factor is derived.

Provincial estimates of the number of families and persons not in families were classified by certain characteristics and used to apply weights to the original sample. Weighting characteristics used were the following:

(a) family size (whether a person not in family or a family of two or more);

(b) sex of unit head;

(c) labour force status of the head (whether paid worker, self-employed - non-farm; self-employed - farm, not in the labour force).

On the basis of the census family definition employed in the Surveys of Consumer Finances, independent estimates of the number of family units by the above characteristics were developed on a provincial basis using special unpublished data from the 1966 Census and various other Statistics Canada sources. These estimates were then reduced to 10 percent of the actual estimate.

A.3 Data Collection

The enumeration procedures used in this survey followed the pattern of earlier income surveys. Information on all household members was recorded on control cards and income questionnaires were left in households for completion by each member 14 years of age and over. Questionnaires were picked up by enumerators a few days later. In cases where income data had not been obtained, a further attempt to solicit this information was made by mail. Data on labour force characteristics of the surveyed population was obtained from the April 1972 Labour Force Survey. The survey reference week for labour force data is the week ending April 22, 1972.

APPENDIX B

CHOICE OF SAMPLES FOR ESTIMATION

B.1 Male Earnings Equations

Of the 25,924 households, 15,379 had male household heads who also satisfied the following conditions.

(a) They were not permanently unable to work.

(b) They did not keep house or attend school full time.

(c) They were not retired or voluntarily idle.

(d) They were not classified as "others" by Statistics Canada.

B.2 Female Earnings Equations

Of the total number of households, 7,609 were used in estimating the female earnings function. They satisfied the following conditions:

(a) They were either married to those satisfying conditions (a) through (d) above or were widowed.

(b) They worked at least part time.

The reason for not including unmarried females will become clear in B.4 below.

B.3 Discriminant Analysis

A smaller sample was used to predict whether the female worked or not. Observations on 2269 females who were married or widowed were drawn at random from the larger sample. Of these, 838 worked, 1431 did not.

B.4 Earnings Profiles

The choice of what set of households to include in the estimation of the annuities is not a minor one. It is clear that the more representative the sample is of the population as a whole, the more useful are the results. In fact, however, there is a tradeoff between the reliability and the generality of the final product. On the one hand, we would like a perfectly representative sample while on the other, it must be recognized that it is not possible to predict with any degree of reliability the life-cycle behaviour of many household heads. With these two principles in mind, the results are based on a sample which includes households which have at least two adults <u>or</u> in the case of single person households where the head is male - and also subject to conditions (a) to (d) in B.1 above on the male member. Thus, it was not considered possible to predict with reliability the lifetime earnings of, for example, unmarried females, males voluntarily idle etc.. While this does restrict the applicability of the results, it does not unduly limit the conclusions which can be drawn from them as the distribution of the annuities is comparable with the distribution of observed incomes and earnings of the same set of households.

APPENDIX C

REGRESSION VARIABLES

(a) Education

Code

(1) Under Grade 5
(2) Grade 5-8
(3) High School not completed
(4) High School completed - vocational program
(5) High School completed - academic program
(6) Non-university some
(7) Non-university completed
(8) University some
(9) University completed - diploma or certificate
(10) University completed - degree

(b) Area of Work

(1) Urban centre with population above 100,000
(2) Urban centre with population of 30,000-99,999
(3) Cities with population of 15,000-29,999
(4) Small urban areas with population 1,000-14,999
(5) Other small towns and rural areas

(c) Class of Worker

(1) Paid employee
(2) Self-employed
(3) Unpaid family worker
Never worked (unemployed)
Not in labour force

(d) Occupational Classification

(1) Managerial occupations
(2) Professional and technical occupations
(3) Clerical occupations
(4) Sales occupations
(5) Service and recreation occupations
(6) Transport and communication occupations
(7) Farmers and farm workers
Loggers and related workers
Fishermen, trappers and hunters
(8) Miners, quarrymen and related workers
Craftsmen, production process and related workers

(9) Labourers not included above
(10) Occupation not stated
(11) Individuals who did not work, look for work or have a job during the reference week

(e) <u>Farm or Non-Farm Earnings</u>

(1) Non-farm
(2) Farm

(f) <u>Experience</u>

(1) less than 1 year
(2) 2, 3 years
(3) 4-6 years
(4) 7-12 years
(5) 13-20 years
(6) 21-30 years
(7) 31-40 years
(8) 41-50 years
(9) 51-60 years
(10) more than 60 years.

APPENDIX D

MINCER SCHOOLING MODEL

Code	Number of Cases	Coefficients
1	17	-.645
2	176	-.127
3	177	-.029
4	25	.103
5	80	.076
6	12	.271
7	20	.389
8	16	.185
9	8	.550
10	45	.251

Dependent variable: log of earnings
Mean of dependent variable: 8.56
Number of observations: 576
Total sum of squares: 1015.8
Explained sum of squares: 20.5

APPENDIX E

UNCERTAINTY

If uncertainty is introduced, the consumer can be considered to maximize the expected value of lifetime utility. In this case, equation (3.1) becomes

$$U = \int_0^{\overline{T}} U[C(t)]e^{-\delta t} R(t)\, dt$$

where R(t) is the probability of being alive in year t. Following the procedure outlined in the text, equation (3.9a) would be replaced by

$$\frac{\dot{C}}{C} = \frac{1}{1-\beta} [r-\delta + \dot{R}/R]$$

where $\dot{R}/R$ is the negative of the conditional probability of dying.

Yaari (1965) shows that under such conditions, the introduction of insurance yields conditions governing optimal consumption behaviour which are identical to the conditions for the certainty case. The difference between the certainty case and the uncertainty case with insurance is that the constants of integration do not coincide.

To the extent that no one ever knows with certainty what their life span will be, the measures being estimated here should best be viewed as the annuities which would define the distribution of economic welfare if individuals had this information.

APPENDIX F

RANDOM NUMBERS

1. The expected value of the variance in the actual data is given by

$$E\ 1/N\ [(Y-\overline{Y})'(Y-\overline{Y})] \tag{F.1}$$

$$= E\ 1/N\ [Y'Y-2Y'\overline{Y}+\overline{Y}'\overline{Y}] \tag{F.2}$$

using the fact that $Y = XB+\varepsilon$ we obtain

$$1/N\ B'X'XB + \sigma^2 - E\ \overline{Y}^2 \tag{F.3}$$

2. The expected value of the variance in the predicted values is given by

$$E\ 1/N\ [(\hat{Y}+\eta-\overline{Y})'(\hat{Y}+\eta-\overline{Y})] \tag{F.4}$$

$$= E\ 1/N\ \hat{Y}'\hat{Y} + E\ 1/N\eta'\eta - EY^2 \tag{F.5}$$

Using the fact that $\hat{Y} = X\hat{B}$, substituting in turn for $\hat{B} = (X'X)^{-1}X'Y$ and $Y = XB+\varepsilon$ and taking the expectation operator through, we obtain

$$1/N\ B'X'XB+K/N\ \sigma^2 + E\ 1/N\ \eta'\eta - E\overline{Y}^2 \tag{F.6}$$

equating (F.3) and (F.6) we obtain

$$(N-K)\sigma^2 = E\eta'\eta$$

Again, let the variance of η_i be V for all i. Thus,

$$V = \frac{(N-K)}{N}\sigma^2$$

An unbiased estimate of σ^2 is $S^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{N-K}$.

Thus, the variance on the random numbers should be

$$V = \frac{N-K}{N}\,\frac{\hat{\varepsilon}\;\hat{\varepsilon}}{N-K} = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{N}$$

APPENDIX G

FAMILY ALLOWANCES REGRESSION

Code	Number of Cases	Coefficients
1	138	-134.3
2	137	-35.9
3	85	38.3
4	43	132.5
5	24	202.0
6	12	384.0
7	8	422.2
8	2	476.5
9	1	-
10	0	-
11	0	-
12	1	914.5

Dependent variable: Family Allowances

Mean of dependent variable: 207.5

Number of observations: 451

Total sum of squares: 11,077,570

Explained sum of squares: 9,051,865

APPENDIX H

Where the utility functional is

$$V = \int_0^T U_t(C_t/F_t)\ e^{-\delta t}\ dt \qquad (H.1)$$

and the instantaneous function is

$$U_t = \frac{A}{\beta}\ (C_t/F_t)^{\beta} \qquad (H.2)$$

the Hamiltonian is given by

$$H = F^{-\beta}C^{\beta}e^{-\delta t} + \lambda S \qquad (H.3)$$

The necessary conditions for a maximum yield a differential equation defining optimal consumption behaviour

$$\dot{C}/C = \frac{-\beta}{1-\beta}\ \dot{F}/F + (r-\delta)/(1-\beta) \qquad (H.4)$$

which when integrated yields a solution

$$c_t^* = \frac{C_o}{F_o^b}\ e^{gt}\ \frac{1}{F_t^{1-b}} \qquad (H.5)$$

where $b = -\beta/(1-\beta)$ and $g = (r-\delta)/(1-\beta)$.

A solution for the annuity is obtained by solving

$$\int_0^T \frac{A}{\beta}(\overline{C}/F_t)^{\beta}e^{-\delta t}\ dt = \int_0^T \frac{A}{\beta}\ c_t^{*\beta}\ e^{-\delta t}\ dt \qquad (H.6)$$

for $\overline{C}$ where c_t^* is defined in (H.5). The solution is thus

$$\overline{C}^{\beta} = [\int_0^T c_t^{*\beta}e^{-\delta t}dt]\ [\int_0^T F_t^{-\beta}e^{-\delta t}\ dt]^{-1} \qquad (H.7)$$

For estimation purposes, this can be simplified by substituting for $c_t^{*\beta}$ using relation (H.5). (H.7) then becomes

$$\bar{C}^{\beta} = [\int_0^T \{\frac{C_o}{F_o^b} e^{gt} F_t^{b-1}\}^{\beta} e^{-\delta t} dt] [\int_0^T F_t^{-\beta} e^{-\delta t} dt]^{-1} \quad \text{(H.8)}$$

We can evaluate $\frac{C_o}{F_o^b}$ by using the fact that lifetime wealth is exhausted by lifetime consumption. i.e.

$$\int_0^T F_t \, C_t/F_t \, e^{-rt} \, dt = W_o \quad \text{(H.9)}$$

By again making use of (H.5), the solution for C_c/F_o^b is

$$\frac{C_o}{F_o^b} = W_o/[\int_0^T F_t^b \, e^{(g-r)t} \, dt] \quad \text{(H.10)}$$

which, when inserted into (H.8) yields a value for the annuity

$$\bar{C} = W_o [\int_0^T F_t^b e^{(g-r)t} dt]^{-1} [\frac{\int_0^T F_t^b \, e^{(g\beta-\delta)t} dt}{\int_0^T F_t^{-\beta} e^{-\delta t} \, dt}] \quad \text{(H.11)}$$

APPENDIX I

The data pertaining to net worth on the Micro-Data Tape 1971 (Census families) are the following: (1) Net income from investment, (2) tenure in place of residence.

We could conceivably scale up the income from investment to yield a stock value. But income from investments reflects such a small percentage of total wealth for an average family that it would be of little use in determining net worth for the vast majority of families.

A more promising approach would be to estimate the capital value of the place of residence since a typical Canadian family holds about 65% of their total wealth in the form of house equity. Unfortunately, the data define only whether the household owns the place of residence or not. No information is available on the equity the household owns.

Given the shortage of useful information, we must estimate the net asset position rather than knowing it with certainty.

APPENDIX J

SENSITIVITY TESTS

In this appendix, we present the results of some sensitivity tests carried out on the computational procedures.

J.1 Random Components

It is possible that different random elements used in the computations could lead to different results. To test for this possibility, the model has been reestimated, starting the random number generator at a different point each time. Table J.1 contains the results for a limited number of the inequality statistics for the annuities using the welfare function U = FU(C/F). As can be seen, the measures are quite insensitive, with differences usually being reflected at the third decimal place.

J.2 Different Samples

To test for the sensitivity of the samples used, several other samples were generated and the inequality statistics estimated on the basis of these. The results are given in Table J.2. Again, these are not significantly different from the results presented in the text.

J.3 Distribution of Random Numbers

The random numbers in the main computations are uniformly distributed. It is possible, if these numbers were distributed in a different manner, that the inequality statistics would not be identical even if the samples were very large and generated subject to the same variance. To test for this, the model has been rerun using normal random numbers. The results are given in Table J.3.

TABLE J.1

	I	II	III	IV
Coefficient of Variation	.5064	.5035	.5148	.5181
Standard Deviation of Logarithms	.6065	.6011	.6186	.6133
Gini Index	.2832	.2804	.2875	.2876
Atkinson's Measure ($\varepsilon = 1$)	.1372	.1349	.1418	.1402

Cumulative Distribution

I	.0268	.0731	.1357	.2115	.2996	.3990	.5132	.6438	.8044	1.0
II	.0273	.0745	.1378	.2137	.3012	.4024	.5164	.6469	.7975	1.0
III	.0257	.0711	.1341	.2102	.2973	.3972	.5099	.6406	.7958	1.0
IV	.0277	.0740	.1345	.2088	.2971	.3972	.5109	.6399	.7917	1.0

Inequality statistics defining the annuity distribution based on different starting points for the random number generator.

TABLE J.2

	I	II	III	IV
Coefficient of Variation	.5064	.5043	.5027	.5342
Standard Deviation of Logarithms	.6065	.6033	.6022	.7060
Gini Index	.2832	.2828	.2807	.2983
Atkinson's Measure ($\varepsilon = 1$)	.1372	.1364	.1350	.1580

Cumulative Distribution

I	.0268	.0731	.1354	.2115	.2996	.3990	.5132	.6438	.8004	1.0
II	.0271	.0731	.1352	.2113	.2999	.4002	.5140	.6448	.7990	1.0
III	.0276	.0745	.1360	.2120	.2997	.4018	.5173	.6483	.7995	1.0
IV	.0249	.0675	.1271	.2006	.2887	.3909	.5043	.6337	.7914	1.0

Inequality statistics defining the annuity distribution based on four different random samples.

TABLE J.3

INEQUALITY STATISTICS FOR ANNUITIES USING U=FU(C/F) AND NORMALLY DISTRIBUTED RANDOM NUMBERS

No. of Observations	659
No. of Weighted Observations	19455
Mean	12332.82
Relative Mean Deviation	.3987
Variance	38258898.0
Coefficient of Vibration	.5015
Standard Deviation of Logs	.6252
Theil's Entropy Measure	.1263
Gini Index	.2797

Decile Shares	.0267	.0482	.0642	.0757	.0874	.1002	.1144	.1288	.1515	.2028
Cumulative	.0267	.0749	.1391	.2148	.3022	.4024	.5169	.6457	.7972	1.0000

(First 9 deciles each form 9.997 percent of total)

Atkinson's Measure (Epsilon = .49)	.0641
Atkinson's Measure (Epsilon = .99)	.1362
Atkinson's Measure (Epsilon =1.49)	.2636
Atkinson's Measure (Epsilon =1.99)	.8366